



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 144864

TO: Jeffrey Parkin
Location: rem/3d39/3c18
Art Unit: 1648
Monday, February 28, 2005

Case Serial Number: 09/319156

From: Deirdre Arnold
Location: Biotech-Chem Library
REM 1A64
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

Please contact me if you encounter any problems with the disk or files.

Note that results are not generally sent via e-mail because the files are often very large.

Please feel free to contact me if you have any questions or would like to amend the search.

Thank you for using STIC services.

Regards,
Deirdre Arnold

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

144864

me

From: Parkin, Jeffrey
Sent: Wednesday, February 09, 2005 9:52 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 09/319,156

Please search SEQ ID NOS.: 6, 9, and 12 from the aforementioned application v. all relevant databases, including interference.

Place results on both PAPER and electronic format (i.e., e-mail).

Provide the first 40 results for each search.

Thanks!

JSP
AU 1648
REM 3D39
2-0908

RECEIVED
FEB 10 2005
STIC-Biotech/ChemLib
(STIC)

http://expoweb1:8001/cgi-bin/expo/GenInfo/snquery.pl?APPL_ID=09319156

STAFF USE ONLY

Searcher: Arnold
Searcher Phone: 2-2533
Date Searcher Picked up: 2/16/05
Date Completed: 2/28/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # 3
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Compu
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:32:54 ; Search time 2725.62 Seconds
(without alignments)
11288.850 Million cell updates/sec

Title: US-09-319-156B-6.
Perfect score: 635
Sequence: 1 cccgtgtatctttaaactctct.....tgaaaaa.....635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	635	100.0	635	6	BD136195 Retrovira
2	635	100.0	635	6	AR344385 Sequence
3	635	100.0	635	6	AX001024 Sequence
4	635	100.0	2004	14	AF127229 Multiple
5	631.8	99.5	167694	9	AC113137 Homo sapi
6	631.8	99.5	176095	2	AC019346 Homo sapi
7	631.8	99.5	210336	2	AC022171 Homo sapi
8	580.2	91.4	148724	2	AC010778 Homo sapi
9	579.2	91.2	8339	9	AL162912 Human DNA
10	571	89.9	136901	9	AC073626 Homo sapi
11	566.4	89.2	176188	9	AC068898 Homo sapi
12	564.6	88.9	89728	9	AL583805 Human DNA
13	560.2	88.2	112405	9	AC072023 Homo sapi
14	559.8	88.2	179114	9	BS000045 Pan trogl
15	556.6	87.7	174019	9	AP001538 Homo sapi
16	556.6	87.7	340000	2	AP001674 Homo sapi
17	551.8	86.9	149755	2	AP001545 Homo sapi
18	551.8	86.9	152980	2	AC090313 Homo sapi
19	551.8	86.9	163803	9	AC093531 Homo sapi

C 20	551.8	86.9	176425	9	AC107075	AC107075 Homo sapi
C 21	551.4	86.8	128468	9	AL139090	AL139090 Human DNA
C 22	551.4	86.8	140756	9	AL139038	AL139038 Human DNA
C 23	551.4	86.8	167366	2	AC021774	AC021774 Homo sapi
C 24	551.4	86.8	183499	2	AL607153	AL607153 Homo sapi
C 25	550.2	86.6	164211	2	AP002790	AP002790 Homo sapi
C 26	550.2	86.6	167843	9	AC064801	AC064801 Homo sapi
C 27	549	86.5	180523	9	CNS01DSH	AL121784 Human chr
C 28	549	86.5	190565	9	AC007374	AC007374 Homo sapi
C 29	549	86.5	203777	9	CNS01DVH	AL135818 Human chr
C 30	547.8	86.3	112544	9	AC109992	AC109992 Homo sapi
C 31	547.6	86.2	192178	2	AC009443	AC009443 Homo sapi
C 32	547.6	86.2	192899	9	CNS05RFS	AL352295 Human chr
C 33	547.4	86.2	172281	9	AC068492	AC068492 Homo sapi
C 34	547.2	86.2	169418	9	AC104163	AC104163 Homo sapi
C 35	546.6	86.1	161054	6	AX478550	AX478550 Sequence
C 36	546.4	86.0	2074	6	AX478550	AX478550 Sequence
C 37	546	86.0	169234	9	CNS06C7Q	AL390799 Human chr
C 38	545	85.8	1329	6	BD136199	BD136199 Retrovira
C 39	545	85.8	1329	6	AR344389	AR344389 Sequence
C 40	545	85.8	1329	6	AX001030	AX001030 Sequence
C 41	545	85.8	99408	9	AL359385	AL359385 Human DNA
C 42	545	85.8	170586	2	AL356126	AL356126 Homo sapi
C 43	543.8	85.6	181166	9	AC087897	AC087897 Homo sapi
C 44	543.4	85.6	88328	9	AL357874	AL357874 Human DNA
C 45	539	84.9	2082	2	AC134379	AC134379 Homo sapi

ALIGNMENTS

RESULT 1	BD136195	635 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD136195	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.			
ACCESSION	BD136195.1	GI:232331140			
VERSION	BD136195.1	GI:232331140			
KEYWORDS	JP 2002509437-A/5.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 635)				
AUTHORS	Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F., Perron,H. and Mandrand,B.				
TITLE	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses				
JOURNAL	Patent: JP 2002509437-A 5 26-MAR-2002;				
COMMENT	BIO MERIEUX				
	OS Unidentified				
	PN JP 2002509437-A/5				
	PD 26-MAR-2002				
	PF 07-JUL-1998 JP 1999508255				
	PR 07-JUL-1997 FR 97/08816				
	PI GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC PI BEDIN				
	PI MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON, PI BERNARD MANDRAND				
	PC C12N15/48,C12Q1/70,C07K14/15,A61K31/70				
	CC Strandedness: Single;				
	CC Topology: Linear;				
	CC Retroviral nucleic material and nucleotide fragments, in CC particular, associated with multiple sclerosis and/or rheumatoid arthritis,				
	CC for				
	CC diagnostic, prophylactic and therapeutic uses FH Key				
	Location/Qualifiers				
	1. .635				
	/organism='Unidentified'.				
FT	source				
FT	Location/Qualifiers				

source		1. .635		/organism="unidentified"		/mol_type="genomic DNA"		/db_xref="taxon:32644"	
ORIGIN									
Query Match		100.0%; Score 635; DB 6; Length 635;							
Best Local Similarity		100.0%; Pred. No. 1.1e-189;							
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60						
Db	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60						
Qy	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120						
Db	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120						
Qy	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180						
Db	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180						
Qy	181	GAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240						
Db	181	GAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240						
Qy	241	CATCAGCCAACTCCGCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC	300						
Db	241	CATCAGCCAACTCCGCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC	300						
Qy	301	AGGACTAGCTGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360						
Db	301	AGGACTAGCTGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360						
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTACACCCGACCAATCAGAGAGCTC	420						
Db	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTACACCCGACCAATCAGAGAGCTC	420						
Qy	421	ACTAAATGCTAATTAGGCAGAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480						
Db	421	ACTAAATGCTAATTAGGCAGAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480						
Qy	481	AGAGCACAGGGAGGGACAAAGATCGGGATATAAACCCAGGCATTTCGAGCCGCAACGG	540						
Db	481	AGAGCACAGGGAGGGACAAAGATCGGGATATAAACCCAGGCATTTCGAGCCGCAACGG	540						
Qy	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600						
Db	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600						
Qy	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635						
Db	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635						
RESULT 2									
AR344385									
LOCUS		Sequence 102 from patent US 6582703.							
DEFINITION		635 bp mRNA linear PAT 17-AUG-2003							
ACCESSION		AR344385							
VERSION		AR344385.1							
KEYWORDS		GI:33740326							
SOURCE		Unknown.							
ORGANISM		Unknown.							
REFERENCE		1 (bases 1 to 635)							
AUTHORS		Perron,H., Besme,F., Bedin,F., Paranhos-Baccala,G., Komarian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.							
TITLE		Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting							
JOURNAL		Patent: US 6582703-A 102 24-JUN-2003;							
FEATURES		Location/Qualifiers							
source		1. .635							
		/organism="unidentified"							

/organism="unknown"		/mol_type="mRNA"	
ORIGIN			
Query Match		100.0%; Score 635; DB 6; Length 635;	
Best Local Similarity		100.0%; Pred. No. 1.1e-189;	
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60
Db	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60
Qy	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120
Db	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120
Qy	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180
Db	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180
Qy	181	GAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240
Db	181	GAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240
Qy	241	CATCAGCCAACTCCGCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC	300
Db	241	CATCAGCCAACTCCGCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC	300
Qy	301	AGGACTAGCTGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360
Db	301	AGGACTAGCTGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTACACCCGACCAATCAGAGAGCTC	420
Db	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTACACCCGACCAATCAGAGAGCTC	420
Qy	421	ACTAAATGCTAATTAGGCAGAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480
Db	421	ACTAAATGCTAATTAGGCAGAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480
Qy	481	AGAGCACAGGGAGGGACAAAGATCGGGATATAAACCCAGGCATTTCGAGCCGCAACGG	540
Db	481	AGAGCACAGGGAGGGACAAAGATCGGGATATAAACCCAGGCATTTCGAGCCGCAACGG	540
Qy	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600
Db	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600
Qy	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635
Db	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635
RESULT 3			
AX001024			
LOCUS		Sequence 6 from Patent WO9902666.	
DEFINITION		635 bp DNA linear PAT 10-MAR-2000	
ACCESSION		AX001024	
VERSION		AX001024.1	
KEYWORDS		GI:7241262	
SOURCE		unidentified	
ORGANISM		unclassified.	
REFERENCE		1 (bases 1 to 635)	
AUTHORS		Ost,C. and Bedin,F.	
TITLE		RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES	
JOURNAL		Patent: WO 9902666-A 6 21-JAN-1999;	
FEATURES		Location/Qualifiers	
source		1. .635	
		/organism="unidentified"	

ORIGIN	/mol_type="unassigned DNA" /db_xref="taxon:32644"	
Query Match	100.0%; Score 635; DB 6; Length 635;	
Best Local Similarity	100.0%; Pred. No. 1.1e-189;	
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 CCTGTATCTTTAACTCCTCTGTTAAGTTGTCTCTCCAGATCAAACTGTAAACTA 60	
Db	1 CCTGTATCTTTAACTCCTCTGTTAAGTTGTCTCTCCAGATCAAACTGTAAACTA 60	
QY	61 CAAATTTGTTCTTCAAAATGGAGCAGACAGATGGAGTCCATGACATAAGATCCACCGTGGACCC 120	
Db	61 CAAATTTGTTCTTCAAAATGGAGCAGACAGATGGAGTCCATGACATAAGATCCACCGTGGACCC 120	
QY	121 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180	
Db	121 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180	
QY	181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGACGATTTAGACGGT 240	
Db	181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGACGATTTAGACGGT 240	
QY	241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGGACTTGAGAGAC 300	
Db	241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGGACTTGAGAGAC 300	
QY	301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCTAGCTGGGAAGGTGACT 360	
Db	301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCTAGCTGGGAAGGTGACT 360	
QY	361 GCATCACCCTCTAAACATGSGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCTC 420	
Db	361 GCATCACCCTCTAAACATGSGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCTC 420	
QY	421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCTG 480	
Db	421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCTG 480	
QY	481 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 540	
Db	481 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 540	
QY	541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600	
Db	541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600	
QY	601 ATTAAATCTTGCACCTGAAAAAAGAAAAA 635	
Db	601 ATTAAATCTTGCACCTGAAAAAAGAAAAA 635	
RESULT 4	2004 bp mRNA linear VRL 11-AUG-1999	
AF127229	Multiple sclerosis associated retrovirus element clone CL6	
LOCUS	pol-env/3'-LTR-like mRNA sequence.	
DEFINITION	AF127229	
ACCESSION	AF127229.1 GI:5726294	
VERSION		
KEYWORDS	Multiple sclerosis associated retrovirus element	
SOURCE	Viruses; Retroviridae.	
ORGANISM	1 (bases 1 to 2004)	
REFERENCE	Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F.,	
AUTHORS	Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E.,	
TITLE	Molecular cloning and characterization of MSRV-related sequences	
JOURNAL	associated with retrovirus-like particles	
MEDLINE	Virology 260 (1), 1-9 (1999)	
PUBMED	99335590	
	10403550	
REFERENCE	2 (bases 1 to 2004)	
AUTHORS	Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F.,	
TITLE	Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E.,	
JOURNAL	Mallet, F., Mandrand, B. and Perron, H.	
MEDLINE	Direct Submission	
PUBMED	Submitted (10-FEB-1999) UMR103 CNRS, bioMerieux, 46, allée	
	D'Italie, Lyon 69007, France	
	Location/Qualifiers	
	1..2004	
	/organism="Multiple sclerosis associated retrovirus	
	element"	
	/mol_type="mRNA"	
	/db_xref="taxon:89382"	
	/clone="CL6"	
	1..2004	
	/note="similar to pol-env and 3'-LTR region"	
misc_feature		
ORIGIN		
Query Match	100.0%; Score 635; DB 14; Length 2004;	
Best Local Similarity	100.0%; Pred. No. 1.3e-189;	
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 CCTGTATCTTTAACTCCTCTGTTAAGTTGTCTCTCCAGATCAAACTGTAAACTA 60	
Db	1370 CCTGTATCTTTAACTCCTCTGTTAAGTTGTCTCTCCAGATCAAACTGTAAACTA 1429	
QY	61 CAAATTTGTTCTTCAAAATGGAGCAGACAGATGGAGTCCATGACATAAGATCCACCGTGGACCC 120	
Db	1430 CAAATTTGTTCTTCAAAATGGAGCAGACAGATGGAGTCCATGACATAAGATCCACCGTGGACCC 1489	
QY	121 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180	
Db	1490 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 1549	
QY	181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGACGATTTAGACGGT 240	
Db	1550 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGACGATTTAGACGGT 1509	
QY	241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGGACTTGAGAGAC 300	
Db	1610 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGGACTTGAGAGAC 1669	
QY	301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCTAGCTGGGAAGGTGACT 360	
Db	1670 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCTAGCTGGGAAGGTGACT 1729	
QY	361 GCATCACCCTCTAAACATGSGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCTC 420	
Db	1730 GCATCACCCTCTAAACATGSGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCTC 1789	
QY	421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCTG 480	
Db	1790 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCTG 1849	
QY	481 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 540	
Db	1850 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 1909	
QY	541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600	
Db	1910 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 1969	
QY	601 ATTAAATCTTGCACCTGAAAAAAGAAAAA 635	
Db	1970 ATTAAATCTTGCACCTGAAAAAAGAAAAA 2004	
RESULT 5	167694 bp DNA linear PRI 25-APR-2002	
AC113137/c	AC113137	
LOCUS	Homo sapiens chromosome 18, clone RP11-622J9, complete sequence..	
DEFINITION	AC113137	
ACCESSION	AC113137.2 GI:19683504	
VERSION		

KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E. 1 (bases 1 to 167694)
TITLE	Homo sapiens chromosome 18, clone RP11-622J9
JOURNAL	Unpublished
AUTHORS	2 (bases 1 to 167694)
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C., Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rietta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 167694)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C., Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rietta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (bases 1 to 167694)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C., Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rietta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	5 (bases 1 to 167694)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C., Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rietta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

FEATURES	Location/Qualifiers
source	1. 167694 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="18" /map="18"
repeat_region	/Clone="RP11-622J9" /Clone_lib="RPC1-11 Human Male BAC" 2566..2593 /rpt_family="AT_rich"
repeat_region	3108..3127 /rpt_family="(TA)n"
repeat_region	3128..3153 /rpt_family="(TG)n"
repeat_region	3154..3208 /rpt_family="(TA)n"
repeat_region	complement(3212..3818) /rpt_family="L1MA9"
repeat_region	4840..5079 /rpt_family="L1ME3"
repeat_region	5234..6028 /rpt_family="L1ME3"
repeat_region	complement(6259..6470) /rpt_family="MIR"
repeat_region	6723..6750 /rpt_family="AT_rich"
repeat_region	6982..7280 /rpt_family="AluY"
repeat_region	7780..7811 /rpt_family="AT_rich"
repeat_region	complement(8120..8302) /rpt_family="L1M4"
repeat_region	complement(8444..8971) /rpt_family="L1ME3A"
repeat_region	complement(8973..9386) /rpt_family="L1ME3A"
repeat_region	complement(9387..9690) /rpt_family="L1MA9"
repeat_region	9898..10437 /rpt_family="MLTIE3"
repeat_region	11218..11555 /rpt_family="L2"
repeat_region	complement(12442..13393) /rpt_family="L1PA16"
repeat_region	13550..13907

TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On March 22, 2002 this sequence version replaced gi:18875226. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L24497 Center clone name: 622_J_9 -----

```

/rpt family="LIME"
14235..15632
/rpt family="LIPB1"
15633..15912
/rpt family="TA)n"
15920..16279
/rpt family="LIPB1"
16335..17100
/rpt family="HAL1b"
17572..17628
/rpt family="LIM4"
17684..18430
/rpt family="LIP4A"
18441..19292
/rpt family="L2"
19293..20623
/rpt family="LIMB3"
complement(20633..22006)
/rpt family="LIP"
22008..22831
/rpt family="LIMB3"
22853..23516
/rpt family="L2"
24420..24468
/rpt family="AT_rich"
24886..24913
/rpt family="AT_rich"
25154..25156
/notes="single clone coverage"
25189..25193
/notes="<30 qual SNGL region"
complement(26144..26248)
/rpt family="MIR"
27732..27794
/rpt family="AT_rich"
27889..27937
/rpt family="TA)n"
28013..28078
/rpt family="TAGA)n"
28704..28743
/rpt family="TAGA)n"
29367..29451
/rpt family="GA-rich"
complement(30946..31096)
31912..31989
/rpt family="MER67C"
32017..32303
/rpt family="MER67C"
complement(32304..34133)

Query Match 99.5%; Score 631.8; DB 9; Length 167694;
Best Local Similarity 99.7%; Pred. No. 2.9e-188;
Matches 633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTGTATCTTTAACTCCTTGTAGTTTCTCTCCAGATCAAACTGTAACACTA 60
Db 123609 CCTGTATCTTTAACTCCTTGTAGTTTCTCTCCAGATCAAACTGTAACACTA 123550

Qy 61 CAAATTCCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 123549 CAAATTCCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 123490

Qy 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGACACCCCTCCGAG 180
Db 123489 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGACACCCCTCCGAG 123430

Qy 181 GAAATCTCACTGCACACCCCTACTATGCCCCAATTCAGCGGGAACGATGAGCGGT 240
Db 123429 GAAATCTCACTGCACACCCCTACTATGCCCCAATTCAGCGGGAACGATGAGCGGT 123370

Qy 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300

```

Db 123369 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 123310

Qy 301 AGGACTAGCTGGATTCTCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGTGACT 360

Db 123309 AGGACTAGCTGGATTCTCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGTGACT 123250

Qy 361 GCATCCACCTCTAAACATGCGGCTTGCAACTTGTAGCTCACACCGGACCAATCAGAGAGCTC 420

Db 123249 GCATCCACCTCTAAACATGCGGCTTGCAACTTGTAGCTCACACCGGACCAATCAGAGAGCTC 123190

Qy 421 ACTAAATGCTTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTCCCTG 480

Db 123189 ACTAAATGCTTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTCCCTG 123130

Qy 481 AGAGCACAGCGGAGGACCAAGGATCGGATATAAACCCAGGCAATTCGAGCCGCAACGG 540

Db 123129 AGAGCACAGCGGAGGACCAAGGATCGGATATAAACCCAGGCAATTCGAGCCGCAACGG 123070

Qy 541 CAACCCCTTTGGGTCCTCTTTGTATGGGCGCTCTGTTTCACTCTATTTCCTCT 600

Db 123069 CAACCCCTTTGGGTCCTCTTTGTATGGGCGCTCTGTTTCACTCTATTTCCTCT 123010

Qy 601 ATTAATCTTGAACATGAAAAAATAAAAAA 635

Db 123009 ATTAATCTTGAACATGAAAAAATAAAAAA 122975

RESULT 6

AC019346 176095 bp DNA linear HTG 25-OCT-2001

LOCUS Homo sapiens chromosome 18 clone RP11-497M7 map 18, WORKING DRAFT

DEFINITION SEQUENCE.

AC019346 GI:16418201

VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 176095)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 18, clone RP11-497M7

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 176095)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F., Boguslavsky,L., Boukhgaiter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenescor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehocsky,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo.A., Wu.X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

TITLE Submitted (02-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL On Oct 25, 2001 this sequence version replaced gi:7230183.

COMMENT All repeats were identified using RepeatMasker:

Smitt, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

1 66468: contig of 66468 bp in length

FEATURES
source
* 66469 66568: gap of unknown length
* 66569 210336: contig of 143768 bp in length.
Location/Qualifiers
1. .210336
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/clone="RP11-407C18"
/clone_lib="RP11 human BAC library 11"
misc_feature
1. .66468
/note="assembly_name:Contig44
clone_end:T7"
misc_feature
66569_210336
/note="assembly_name:Contig45
clone_end:SP6"

ORIGIN

Query Match 99.5%; Score 631.8; DB 2; Length 210336;
Best Local Similarity 99.7%; Pred. No. 3e-188;
Matches 633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCTGTATCTTTAACTCTCTGTTAAAGTTGCTCTTCAGAAATCAAAATCTGTAAACTA 60
Db 58176 CCTGTATCTTTAACTCTCTGTTAAAGTTGCTCTTCAGAAATCAAAATCTGTAAACTA 58235
QY 61 CAAATTTCTTCAATGAGCAGCAGATGGATCCATGACATGAAGTCCACCGTGACCC 120
Db 58236 CAAATTTCTTCAATGAGCAGCAGATGGATCCATGACATGAAGTCCACCGTGACCC 58295
QY 121 CTGACCGGCTCTAGCCATCTCTCGATGTTAAATGAATGAAGCGACCCCTCCCGAG 180
Db 58296 CTGACCGGCTCTAGCCATCTCTCGATGTTAAATGAATGAAGCGACCCCTCCCGAG 58355
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAAGCATAGACGGT 240
Db 58356 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAAGCATAGACGGT 58415
QY 241 CATCAGCAACCTCCCAACAGCACTTGGTCTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 58416 CATCAGCAACCTCCCAACAGCACTTGGTCTTCTGTTGAGAGGGGAGCTGAGAGAC 58475
QY 301 AGGACTAGCTGGATTTCTTAGGCAACAGAAATCCTTAAGCCTAGCTGGGAAGGTGACT 360
Db 58476 AGGACTAGCTGGATTTCTTAGGCAACAGAAATCCTTAAGCCTAGCTGGGAAGGTGACT 58535
QY 361 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 58536 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 58595
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG 480
Db 58596 ACTAAATGCTAATAGGCAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG 58655
QY 481 AGAGCAGCGGAGGAGCAAGATCGGATATAAACCCAGGATTCGAGCCGCGCAACGG 540
Db 58656 AGAGCAGCGGAGGAGCAAGATCGGATATAAACCCAGGATTCGAGCCGCGCAACGG 58715
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTCTTTTCACTCTATTTCACCT 600
Db 58716 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTCTTTTCACTCTATTTCACCT 58775
QY 601 ATTAATCTTGCACACTGAAAAAAGAAAAA 635
Db 58776 ATTAATCTTGCACACTGAAAAAAGAAAAA 58810

RESULT 8

AC010778/c 148724 bp DNA linear HTG 30-MAR-2000
LOCUS AC010778 Homo sapiens clone RP11-2N15, WORKING DRAFT SEQUENCE, 8 unordered
DEFINITION pieces.
AC010778
AC010778

VERSION
KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC010778.3 GI:7341824
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 148724)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-2N15
Unpublished
2 (bases 1 to 148724)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6087973.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2677
Center clone name: 2.N.15
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 140899 bases at least Q40
Consensus quality: 144351 bases at least Q30
Consensus quality: 145949 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 148024; sum-of-contigs
Quality coverage: 3.9 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 2462: contig of 2462 bp in length
* 2463 2562: gap of 100 bp
* 2563 7948: contig of 5386 bp in length
* 7949 8048: gap of 100 bp
* 8049 15800: contig of 7752 bp in length
* 15801 15900: gap of 100 bp
* 15901 25059: contig of 9159 bp in length
* 25060 25159: gap of 100 bp
* 25160 41906: contig of 16747 bp in length
* 41907 42006: gap of 100 bp
* 42007 63341: contig of 21335 bp in length
* 63342 98587: contig of 35146 bp in length
* 98588 98587: contig of 35146 bp in length

QY	601	ATTAAATCTTCAACTGAAAAAAGAAAAA 633
Db	39671	ATTAAATCTTCAACTGAAAAAAGAAAAA 39639
RESULT 9		
LOCUS	AL162912	Human DNA sequence from clone RP6-198C4 on chromosome Xq13.3-21.2, complete sequence.
DEFINITION	AL162912	AL162912 AL121824
ACCESSION	AL162912.1	GI:7406722
VERSION	HTG.	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 8339)	
AUTHORS	Pearce,A.	
TITLE	Direct Submission	
JOURNAL	Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk	
COMMENT	Clone requests: clonerequest@sanger.ac.uk	
	During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.	
	The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at	
	http://www.sanger.ac.uk/Projects/C_elegans/wormpep	
	This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at	
	http://www.sanger.ac.uk/HGP/ChrX	
	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP6-198C4 is from the library RPCI-6 constructed by the group of Pieter de Jong. For further details see	
	http://www.chori.org/bacpac/home.htm	
	VECTOR: pPAC4	
FEATURES	Location/Qualifiers	
source	1..8339	/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosomes="X"
		/map="ql3.3-21.2"
		/clone="RP6-198C4"
		/clone_lib="RPCI-6"
repeat_region	1..5641	/notes="HERV17 repeat: matches 368. .5930 of consensus"
repeat_region	5635..5673	/note="HERV17 repeat: matches 7419. .7457 of consensus"
repeat_region	5664..5955	/notes="HERV17 repeat: matches 8244. .8523 of consensus"
repeat_region	5956..6136	/notes="LTR17 repeat: matches 1. .239 of consensus"
repeat_region	6195..6280	/notes="LTR17 repeat: matches 239. .324 of consensus"
ORIGIN		
	Query Match	91.4%; Score 580.2; DB 2; Length 148724;
	Best Local Similarity	94.8%; Pred. No. 6.6e-172; Indels 0; Gaps 0;
	Matches 600; Conservative	0; Mismatches 33;
QY	1	CCCTGTATCTTTAACTCTCTGTTAGTTTGTCTTCAGAAATCAAACTGTAAACTA 60
Db	40271	CCCTGTATCTTTAACTCTCTGTTAGTTTGTCTTCAGAAATCAAACTGTAAACTA 40212
QY	61	CAAATTTGTTCTCAATGAGCACCAGATGGATCCATGATCAAGTCCACCGTGAGCC 120
Db	40211	CAAATTTGTTCTCAATGAGCACCAGATGGATCCATGATCAAGTCCACCGTGAGTC 40152
QY	121	CTGGACCGGCTGTAGCCATGCTCCGATGTTAATGACATTTGAAGGCACCCCTCCGAG 180
Db	40151	CTGGACCGGCTGTAGCCATGCTCCGATGTTAATGACATTTGAAGGCACCCCTCCGAG 40092
QY	181	GAAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCATAGACGGT 240
Db	40091	GAAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCATAGACGAGT 40032
QY	241	CATCAGCCACCTCCCAACGACCTGGTTTCTGTTGAGGGGGGACTGAGAGAC 300
Db	40031	CATCAGCCACCTCCCAACGACCTGGGTGTTCTGTTGAGAGGGGGTACTGAGAGAC 39972
QY	301	AGGACTAGCTGGATTTCTTAGGCCAACGAAATCCCTAAGCTAGCTGGGAAGGTGACT 360
Db	39971	AGGACTAGCTGGATTTCTTAGGCCAACGAAATCCCTAAGCTAGCTGGGAAGGTGACC 39912
QY	361	GCATCCACCTCTTAAACATGGGGTTGGCACTTAGCTTCACACCGGACCAATCAGAGCTC 420
Db	39911	ACACCCACCTTTAAACATGGGGTTGGCACTTAGCTTCACACCGGACCAATCAGAGCTC 39852
QY	421	ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 480
Db	39851	ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 39792
QY	481	AGAGCAGCGGGGAGGACAAAGGATCGGGATATAAACCAGGCAATTCGAGCCGGCAACGG 540
Db	39791	AGAGCAGCGTGGAGGACAAAGGATCGGGATATAAACCAGGCAATTCGAGCCGGCAACAG 39732
QY	541	CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
Db	39731	CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 39672

AL162912 8339 bp DNA linear PRI 22-NOV-2001
Human DNA sequence from clone RP6-198C4 on chromosome Xq13.3-21.2,
complete sequence.
AL162912 AL121824
AL162912.1 GI:7406722
HTG.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 8339)
Pearce,A.
Direct Submission
Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
Clonerequest@sanger.ac.uk
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. RP6-198C4 is from the
library RPCI-6 constructed by the group of Pieter de Jong. For
further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pPAC4
IMPORTANT: This sequence is not the entire insert of clone
RP6-198C4. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
Location/Qualifiers
FEATURES
source
1..8339
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="q13.3-21.2"
/clone="RP6-198C4"
/clone_lib="RPCI-6"
1..5641
/note="HERV17 repeat: matches 368..5930 of consensus"
repeat_region
5635..5673
/note="HERV17 repeat: matches 7419..7457 of consensus"
repeat_region
5664..5955
/note="HERV17 repeat: matches 8244..8523 of consensus"
repeat_region
5956..6136
/note="L1717 repeat: matches 1..239 of consensus"
repeat_region
6195..6280
/note="L1717 repeat: matches 239..324 of consensus"
ORIGIN

```

Query Match          91.2%; Score 579.2; DB 9; Length 8339;
Best Local Similarity 94.8%; Pred. No. 8.4e-172;
Matches 599; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTGTTAAAGTTTGCTCTTCAGAAATCAAACTGTAATAACTA 60
Db CCTGTATCTTTAACTCTCTGTTAAAGTTTGCTCTTCAGAAATCAAACTGTAATAACTA 5723

QY 61 CAAATGTTCTTCAATGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db CAAATGTTCTTCAATGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGATCC 5783

QY 121 CTGACCGGCTCTAGCCATCTCCGATGTTAAATGACATGAAGCCACCCCTCCCGAG 180
Db CTGACCGGCTCTAGCCATCTCCGATGTTAAATGACATGAAGCCACCCCTCCCGAG 5843

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACGGT 240
Db GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACGAGT 5903

QY 241 CATCAGCAACCTCCCAACAGCAGTCTGGTTTTCCTGTTGAGAGGGGAGCTGAGAGAC 300
Db CATCAGCAACCTCCCAACAGCAGTCTGGTTTTCCTGTTGAGAGGGGAGCTGAGAGAC 5963

QY 301 AGGACTAGCTGGATTTCTTAGGCAAGAGATCCTTAAGCCTAGCTGGGAAGGTGACT 360
Db AGGACTAGCTGGATTTCTTAGGCAAGAGATCCTTAAGCCTAGCTGGGAAGGTGACT 6023

QY 361 GCATCCACCTCTAAACATGSGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGTTC 420
Db GCATCCACCTCTAAACATGSGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGTTC 6083

QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTCCCTG 480
Db ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTCCCTG 6143

QY 481 AGAGCAGCGGAGGAGCAGAGATCGGATATAAACCAGGATTCGAGCCCGGCAACGG 540
Db AGAGCAGCGGAGGAGCAGAGATCGGATATAAACCAGGATTCGAGCCCGGCAACAG 6203

QY 541 CAACCCCTCTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db CAACCCCTCTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 6263

QY 601 ATTAATCTTGCACGTAAGAAAAAAGAAAAA 632
Db ATTAATCTTGCACGTAAGAAAAAAGAAAAA 6295

RESULT 10
AC073626/c 136901 bp DNA linear PRI 30-JAN-2004
LOCUS Homo sapiens BAC clone RP11-95P9 from 7, complete sequence.
DEFINITION AC073626
ACCESSION AC073626.7 GI:12863221
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 136901)
Hillier, L.W., Fulton, R.S., Fulton, L.A., Graves, T.A., Pepin, K.H.,
Wagner-McPherson, C., Layman, D., Maas, J., Jaeger, S., Walker, R.,
Wylie, K., Sekhon, M., Becker, M.C., O'Laughlin, M.D., Schaller, M.E.,
Fewell, G.A., Delehaunty, K.D., Miner, T.B., Nash, W.E., Cordes, M.,
Du, H., Sun, H., Edwards, J., Bradshaw-Cordum, H., Ali, J., Andrews, S.,
Isak, A., Vanbrunt, A., Nguyen, C., Du, F., Lamar, B., Courtney, L.,
Kalicki, J., Ozersky, P., Bielecki, L., Scott, K., Holmes, A.,
Harkins, R., Harris, A., Strong, C.M., Hou, S., Tomlinson, C.,
Dauphin-Kohlberg, S., Kozlowicz-Reilly, A., Leonard, S., Rohlfing, T.,
Rock, S.M., Tin-Wollam, A.M., Abbott, A., Minx, P., Maupin, R.,
Stromwatt, C., Latreille, P., Miller, N., Johnson, D., Murray, J.,
Weesener, J.P., Wendt, M.C., Yang, S.P., Schultz, B.R., Wallis, J.W.,

```

Spith, J., Bieri, T.A., Nelson, J.O., Berkowicz, N., Wohldmann, P.E., Cook, L.L., Hickenbotham, M.T., Eldred, J., Williams, D., Bedell, J.A., Mardis, E.R., Clifton, S.W., Chissoe, S.L., Marra, M.A., Raymond, C., Haugen, E., Gillett, W., Zhou, Y., James, R., Phelps, K., Iadonato, S., Bubb, K., Simma, E., Levy, R., Clendinning, J., Kaul, R., Kent, W.J., Furey, T.S., Baerescu, R.A., Brent, M.R., Keibler, E., Flicek, P., Bork, P., Suyama, M., Bailey, J.A., Portnoy, M.E., Torrens, D., Chinwalla, A.T., Gish, W.R., Eddy, S.R., McPherson, J.D., Olson, M.V., Eichler, E.E., Green, E.D., Waterston, R.H. and Wilson, R.K.

The DNA sequence of human chromosome 7
Nature 424 (6945), 157-164 (2003)
22737999

MEDLINE
PUBMED
12853948
2 (bases 1 to 136901)
Bauer, H., Haakenson, B. and Nguyen, C.
The sequence of Homo sapiens BAC clone RP11-95P9

TITLE
Unpublished (2001)
JOURNAL
3 (bases 1 to 136901)
Waterston, R.H.
Direct Submission

TITLE
Submitted (27-JUN-2000) Genome Sequencing Center, Washington
JOURNAL
MO 63108, USA
4 (bases 1 to 136901)
Waterston, R.H.

REFERENCE
AUTHORS
TITLE
Submitted (16-FEB-2001) Genome Sequencing Center, Washington
JOURNAL
MO 63108, USA
5 (bases 1 to 136901)
Waterston, R.H.

REFERENCE
AUTHORS
TITLE
Submitted (09-MAY-2001) Department of Genetics, Washington
JOURNAL
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 136901)
Waterston, R.H.

REFERENCE
AUTHORS
TITLE
Submitted (29-APR-2003) Department of Genetics, Washington
JOURNAL
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 136901)
Wilson, R.

REFERENCE
AUTHORS
TITLE
Submitted (30-JAN-2004) Department of Genetics, Washington
JOURNAL
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 16, 2001 this sequence version replaced gi:11597125.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_NH0095P09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is CTD-2023N18, 200 bp overlap the
clone sequenced to the right is CTA-250D13, 200 bp overlap. Actual
start of this clone is at base position 137066 of CTD-2023N18
actual end is at base position 26937 of CTA-250D13.

repeat_region

```

QY 541 CACCCCTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTCTACTCT 600
Db 95137 CACCCCTTGGTCCCTCCCTCAATTTGTATGGAGCTCTGTTTCACTCTATTCTACTCT 95078
QY 601 ATTAATCTTGGCACTGAAAAAATAAAAAA 635
Db 95077 ATTAATCTTGGCACTGCAAAAAATAAAAAA 95043

RESULT 11
AC068898/c 176188 bp DNA linear PRI 29-AUG-2002
DEFINITION Homo sapiens chromosome 10 clone RP11-534L6, complete sequence.
ACCESSION AC068898
VERSION AC068898.8 GI:22539070
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 176188)
Smith, D.R.
Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
Unpublished
2 (bases 1 to 176188)
Smith, D.R.
Direct Submission
Submitted (11-MAY-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
3 (bases 1 to 176188)
Smith, D.R.
Direct Submission
Submitted (29-AUG-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Aug 29, 2002 this sequence version replaced gi:14625082.
FEATURES
source
1..176188
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-534L6"
/clone_lib="RPCI-11"

ORIGIN
Query Match 89.2%; Score 566.4; DB 9; Length 176188;
Best Local Similarity 93.5%; Pred. No. 1.6e-167;
Matches 591; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
Db 57896 CCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 57837

QY 61 CAAATTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 57836 CAAATTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 57777

QY 121 CTGACCGGCTGTAGCCCATGCTCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 180
Db 57776 CTAGACTGGCTGTAGCCCATGCTCGATGTTTAATGACATTTGAAGTCACTCTCTCTGAG 57717

QY 181 GAAATCTCAATGACACACCTTACTATGCCCCAATTCAGGGGAGCAGTTAGAGCGGT 240
Db 57716 GAAATCTCAATGACACACCTTACTATGCCCCAATTCAGGGGAGCAGTTAGAGTGGT 57657

QY 241 CATCAGCCACCTCCCAACAGCACTTGGTTTTCGTGTTGAGAGGGGAGCTGAGAGAC 300
Db 57656 CATCAGCCACCTCCCAACAGCACTTGGTTTTCGTGTTGAGAGGGGAGCTGAGAGAC 57597

QY 301 AGGACTAGCTGGATTCTCTAGGCCCAACGAAGAAATCCCTAAGCCCTAGCTGGGAAGTGACT 360
Db 57596 AGGACTAGCTGGATTCTCTAGGCCGATTAAAGAAATCCCTAAGCCCTAGCTGGGAAGTGACT 57537

```

```

QY 361 GCATCCACCTCTAAACATGGGCTTTGCAACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 57536 GCATCCACCTTTAAACACGGGCTTTGCAACTTAGCTCAACCCGACCAATCAGAGAGCTC 57477
QY 421 ACTAAATGCTAATTAGGCAAAAATAGAGGTAAGAGTAAGAAATAGCCATCATCTATTGCTCG 480
Db 57476 ACTAAATGCTAATTAGGCAAAAACAGGAGGTAAGAGTAAGAAATAGCCATCATCTATTGCTCG 57417
QY 481 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCGCGCAACGG 540
Db 57416 AGAGCACAGTGGGAGGACAGGATTCGATATAAACCCAGGATTCGAGCGCGCAACGG 57357
QY 541 CAAACCCCTTTGGGTCCTCCCTCTTTGTATGGCGCTCTGTTTCACTCTATTCTACTCT 600
Db 57356 CAAACCCCTTTGGGTCCTCCCTCTTTGTATGGGAGCTCTGTTTCACTCTATTCTACTCT 57297
QY 601 ATTAATCTTGGCACTGAAAAAATAAAAAA 632
Db 57296 ATTAATCTTGGCACTGAAAAAATAAAAAA 57265

RESULT 12
AL583805 89728 bp DNA linear PRI 29-SEP-2001
LOCUS Human DNA sequence from clone RP11-134K1 on chromosome 9, complete
DEFINITION sequence.
ACCESSION AL583805
VERSION AL583805.7 GI:15865009
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 89728)
Clark, G.
Direct Submission
Submitted (29-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Oct 2, 2001 this sequence version replaced gi:14702155.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-134K1 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-134K1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP11-134K1 is at 89728 in this
sequence. The true right end of clone RP11-60C15 is at 2000 in this
sequence.

```

```

FEATURES
  source      Location/Qualifiers
1. .89728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-134K1"
/clone_lib="RPC1-11.1"

ORIGIN
Query Match      88.9%; Score 564.6; DB 9; Length 89728;
Best Local Similarity 94.2%; Pred. No. 5.4e-167;
Matches 598; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
DB 83434 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCAGAAATCGAAGCTGTAA 83490

QY 61 CAAATTTGTTCTTCAATGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
DB 83491 TATATGGTTCCTTCAATGAGCACCAGATGGAGTCCATGACTTAAGATCTACCGCGGACCC 83550

QY 121 CTGGACCGGCTCTAGCCATGCTCCGATGTTAAATGATTAAGAGGACACCCCTCCCGAG 180
DB 83551 CTGGACCGGCTCTAGCCATGCTCCGATGTTAAATGATTAAGAGGACACCCCTCCGAG 83610

QY 181 GAAATCTCACTGACCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACCGGT 240
DB 83611 GAAATCTCACTGACCAACCACTACTATGCCCAATTCAGCGGGAAGCAGTTAGACCGGT 83670

QY 241 CATCAGCCAACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGAC 300
DB 83671 TGTGAGCCAACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGAC 83730

QY 301 AGGACTAGCTGGATTTCCTAGCCCAACGAAGATCCCTAAGCCTAGCTGGGAAGTGACT 360
DB 83731 AGGACTAGCTGGATTTCCTAGCGGAATAAAGATCCCTAAGCCTAGCTGGGAAGTGACT 83790

QY 361 GCATCCACTCTTAACATGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 420
DB 83791 GCATCCACTCTTAACACGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 83850

QY 421 ACTAAATCTTAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
DB 83851 ACTAAATCTTAATAGGCAAAACAGGAGTAAAGAAATAGCCCAATCATCTATTGCGCTG 83910

QY 481 AGAGCACAGCGGAGGACAAGATCGGATATAAACCCAGGATTCGAGCGCGCAACGG 540
DB 83911 AGAGCACAGCGGAGGACAAGATCAGGATATAAACCCAGGATTCGAGCGCGCAACAG 83970

QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 83971 CAACCCCTTTGGTCCCTCCCTTTGATGGAGCTCTGTTTCACTCTATTTCACCTCT 84030

QY 601 ATTAATCTTGAACCTGAAAAAATAAAAAAAAAAAAA 635
DB 84031 ATTAATCTTGAACCTGAAAAAATAAAAAAAAAAAAA 84065

RESULT 13
AC072023
LOCUS      Homo sapiens 112405 bp DNA linear PRI 16-JAN-2003
DEFINITION Homo sapiens 3 BAC RP11-3J2 (Roswell Park Cancer Institute Human
AC072023 BAC Library) complete sequence.
ACCESSION AC072023
VERSION   AC072023.9 GI:27764634
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 112405)
AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

```

Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,K.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homai,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Lousegod,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,N., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vaequez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (07-JUN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (22-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (16-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 16, 2003 this sequence version replaced gi:21539113.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

FEATURES

source Location/Qualifiers
1. .112405
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-3J2"
1. .2000
/notes="overlaps bases 38072..40071 of clone AC125608"
/function="clone overlap"
complement(1..261)
/rpt_family="L1M4"
complement(431..913)
/rpt_family="L1M6"
complement(914..1221)
/rpt_family="AluSg"
complement(1222..1808)
/rpt_family="L1M6"
complement(2045..2198)
/rpt_family="L1M6"
2454..2811
/rpt_family="THE1C"
2813..3092
/rpt_family="MLT1I"
3194..3870
/rpt_family="L1PA3"
3871..3891
/rpt_family="AT_rich"
complement(4297..4671)
/rpt_family="WSTD"
4714..4913
/rpt_family="MLT1I"
5632..5670
/rpt_family="AT_rich"
complement(6337..6630)
/rpt_family="AluSx"

misc_feature
1. .2000
/notes="overlaps bases 38072..40071 of clone AC125608"

repeat_region
complement(1..261)
/rpt_family="L1M4"
complement(431..913)
/rpt_family="L1M6"
complement(914..1221)
/rpt_family="AluSg"
complement(1222..1808)
/rpt_family="L1M6"
complement(2045..2198)
/rpt_family="L1M6"
2454..2811
/rpt_family="THE1C"
2813..3092
/rpt_family="MLT1I"
3194..3870
/rpt_family="L1PA3"
3871..3891
/rpt_family="AT_rich"
complement(4297..4671)
/rpt_family="WSTD"
4714..4913
/rpt_family="MLT1I"
5632..5670
/rpt_family="AT_rich"
complement(6337..6630)
/rpt_family="AluSx"

repeat_region 7063..7093
/rpt_family="TTTTA)n"
complement(7094..7374)
/rpt_family="AluSp"
complement(7386..7695)
/rpt_family="AluSx"
7910..8035
/rpt_family="MIR"
8824..10387
/rpt_family="L1PA16"
10388..10412
/rpt_family="AT_rich"
10637..10675
/rpt_family="AT_rich"
10754..10818
/rpt_family="L1PB1"
complement(10830..10905)
/rpt_family="L1PA16"
complement(10956..11184)
/rpt_family="L1"
11295..11658
/rpt_family="L2"
complement(12489..12812)
/rpt_family="L1PA15"
complement(12819..13533)
/rpt_family="L1"
13595..13619
/rpt_family="AT_rich"
14620..14933
/rpt_family="AluSg"
complement(15183..15254)
/rpt_family="L2"
complement(16463..17175)
/rpt_family="L1P"
17173..18635
/rpt_family="L1PA7"
18793..19462
/rpt_family="L1M4"
19461..20130
/rpt_family="L1ME"
20776..21073
/rpt_family="AluSx"

Query Match 88.2%; Score 560.2; DB 9; Length 112405;
Best Local Similarity 93.8%; Pred. No. 1.4e-165;
Matches 594; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTTGTAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60
DB 67581 CCTGTATCTTTAACTCTCTTGTAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 67640
QY 61 CAAATGTTCTTCAATGGAGACGAGTGGAGTCCATCACTAAGATCCACCGTGACCC 120
DB 67641 CAAATGGGTCTTCAATGGAGACGAGTGGAGTCCATCACTAAGATCCACCGTGACCC 67700
QY 121 CTGGACCGGCTCTAGCCGATGCTCCGATGTTAATGACATTCGAAGCAGCCCTCCCGAG 180
DB 67701 CTGGACTGGCTGTAGCCGATGATCCGATGTTAATGACATTCGAAGCAGCCCTCCCGAG 67760
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGCGGT 240
DB 67761 GAAATCTCAACTGCCCAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGCGGT 67820
QY 241 CATCAGCCAACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGAGTGGAGAC 300
DB 67821 CATTCGCCAACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGAGTGGAGAC 67880
QY 301 AGGACTAGCTGGATTTCTTAGCCCAAGAGATCCCTAGCCCTAGCTAGCTAGCTAGCTAGCT 360
DB 67881 AGGACTAGCTGGATTTCTTAGCCCAAGAGATCCCTAGCCCTAGCTAGCTAGCTAGCTAGCT 67940
QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTACCTAGCTCAGACCCGACCAATCAGAGAGCTC 420

```
Db 67941 GCATCCACCTTTAAACACGGGCGTCGCAACTTAGCTCACACCAACCAATCAGAGAGCTC 68000
Qy 421 ACTAAATGCTTAATTAGGCAAAATAGGAGTAAGAAATAGCAATCATCTATTGCTG 480
Db 68001 ACTAAATGCTTAATTAGGCAAAACAGAGGTAAT- TAATAGCAATCATCTATTGCTG 68059
Qy 481 AGAGCACAGCGGAGGACAGAGTTCGGATATAAACCCAGGCATTCGAGCGGCAACGG 540
Db 68060 AGAGCACAGCGGAGGACAGAGTTCGGATATATTAACAGGCATTCGAGCGGCAATGG 68119
Qy 541 CAACCCCTTTGGGTCCTCCCTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 68120 CAACCCCTTTGGGTCCTCCCTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 68179
Qy 601 ATTAATCTTGCACCTGAAAAAATAAAAA 633
Db 68180 ATTAATCTTGCACCTGCAAAAAAATAAAAA 68212

RESULT 14
BS000045/c
LOCUS
DEFINITION
BS000045 179114 bp DNA linear PRI 12-JUN-2004
Pan troglodytes chromosome 22 clone:RP43-179P23, map 22, complete
sequences.
ACCESSION
BS000045 BA000046
VERSION
BS000045.1 GI:37537312
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 The International Chimpanzee Chromosome 22 Consortium.
DNA sequence and comparative analysis of chimpanzee chromosome 22
Nature 429, 382-388 (2004)
2 (bases 1 to 179114)
Hattori, M., Toyoda, A., Watanabe, H., Taylor, T.D., Kuroki, Y.,
Fujiyama, A. and Sakaki, Y.
Direct Submission
Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chinese National Human Genome Center at Shanghai, Shanghai, China;
*GFP, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KRIIBB Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.
----- RIKEN Genomic Sciences Center
Center: RIKEN Genomic Sciences Center
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: The Chimpanzee Chromosome 22 Sequencing Project
Center clone name: RP43-179P23
----- Summary Statistics
Sequencing vector: pUC18, pUC13, pTZ19; 100% of reads Chemistry:
Phrap; version 0.990329
Dye-terminator Big Dye and ET; 100% of reads Assembly program:
Consensus quality: 178,810 bases at least Q40
Consensus quality: 263 bases at least Q30
Consensus quality: 35 bases at least Q20
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
```

```
30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at one
plasmid
subclone or more than one M13 subclone;
and the assembly was confirmed by restriction digest.
-----
Source information:
The RPCI-43 chimpanzee BAC library was prepared from DNA isolated
from the blood of a single male chimpanzee using published
protocols (Osoegawa, K. et al. Genomics 52:1-8). The DNA from the
chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
characteristics are described at
http://www.chori.org/bacpac/mchimp43.htm.
The clone may be obtained from Pieter J. de Jong and coworkers
(http://www.chori.org/bacpac).
VECTOR: pBACe3.6
The CHORI-251 chimpanzee BAC library was prepared from DNA isolated
from the blood of a single male chimpanzee using published
protocols (Osoegawa, K. et al. Genomics 52:1-8). The DNA from the
chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
characteristics are described at
http://www.chori.org/bacpac/chimpanzee251.htm.
The clone may be obtained from Pieter J. de Jong and coworkers
(http://www.chori.org/bacpac).
VECTOR: pTARBAC2.1
The PTB1 chimpanzee BAC library was prepared from DNA isolated from
cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKSI145
The PTF22 chimpanzee Fosmid library was prepared from DNA isolated
from cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKSI143
-----
Sequence Quality Assessment:
This entry has been annotated with sequence
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in
10,000 bp.
-----
Neighboring clones: RP43-014B20(left) and PTB-042H12(right).
-----
FEATURES
source
1. 179114
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="RP43-179P23"
/clone_lib="RPCI-43 chimpanzee BAC"
-----
ORIGIN
Query Match 88.2%; Score 559.8; DB 9; Length 179114;
Best Local Similarity 94.0%; Pred. No. 2e-165;
Matches 582; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Qy 1 CCCTGTATCTTTAACTCCCTTGTAGTTGCTCTCTCCAGATCAAAACTGTAACCTA 60
Db 169185 CCCTGTATCTTTAACTCCCTTGTAGTTGCTCTCTCCAGATCAAAACTGTAACCTA 169126
Qy 61 CAAATTGTTCTTCAATGGAGCCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 169125 CAAATAGTTCTTCAATGGAGCCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 169066
```



```
QY 121 CTGAGCCGGCTGTAGCCCTGCTCCGATGTTAATGACATTGAAGGACACCCCTCCCGAG 180
|||
Db 169065 CTAGACTGGCTGTAGCCCTGCTCCGATGTTAATGACATTGAAGGACACCCCTCCCGAG 169006
|||
QY 181 GAAATCTCAACTGCAACAACCCCTACTATGCCCAATTCAGCGGGAGCAGTTAGAGCGGT 240
|||
Db 169005 GAAATCTCAACTGCAACAACCCCTACTATGCCCAATTCAGCGGGAGCAGTTAGAGCGGT 168946
|||
QY 241 CATGAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 300
|||
Db 168945 CATGAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 168886
|||
QY 301 AGGACTAGCTGATTTCTTAGGCCCAACGAAGAATCCCTAAGCTAGCTGGGAAGGTGACT 360
|||
Db 168885 AGGACTAGCTGATTTCTTAGGCCCAACGAAGAATCCCTAAGCTAGCTGGGAAGGTGACT 168826
|||
QY 361 GCATCCACCTCTAAACATGCGGCTTGCAACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
|||
Db 168825 GCATCCACCTCTAAACATGCGGCTTGCAACTTAGCTCAACCCGACCAATCAGAGAGCTC 168766
|||
QY 421 ACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
|||
Db 168765 ACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCGCTG 168706
|||
QY 481 AGAGCACAGCGGAGGACCAAGGATCGGATATAAACCCAGGCATTTCAGAGCGGCAACGG 540
|||
Db 168705 AGAGCACAGCGGAGGACCAAGGATCGGATATAAACCCAGGCATTTCAGAGCGGCAACGG 168646
|||
QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
|||
Db 168645 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 168586
|||
QY 601 ATTAATCTTGAACCTGAA 619
|||
Db 168585 ACTAAATCTTGAACCTGCA 168567
|||
```

```
RESULT 15
AP001538/c
LOCUS
DEFINITION
Homo sapiens genomic DNA, chromosome 21q21.1-q21.2 clone:B853K11,
LL56-APP region, complete sequence.
ACCESSION
AP001538
VERSION
AP001538.1 GI:7328982
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 174019)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Watanabe,H., Yada,T.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 174,019 genomic DNA of 21q21.1-q21.2
Published Only in DataBase (2000)
2 (bases 1 to 174019)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (23-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ. 1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp)
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924
```

```
FEATURES
source
1. 174019
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q21.1-q21.2"
/clone="B853K11"
```

ORIGIN

```
Query Match
Best Local Similarity 87.7%; Score 556.6; DB 9; Length 174019;
Matches 580; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1 CCCTGTATCTTAAACCTCCCTTGTAAAGTTTCTCTTCCAGAAATCAAAACCTGTAACCTA 60
Db 36710 CCCTGTATCTTAAACCTCCCTTGTAAAGTTTCTCTTCCAGAAATCAAAACCTGTAACCTA 36651
|||
QY 61 CAAATGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
|||
Db 36650 CAAATGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 36591
|||
QY 121 CTGACCGCGCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGACACCCCTCCCGAG 180
|||
Db 36590 CTGACCTGGCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGGACACCCCTCCCGAG 36531
|||
QY 181 GAAATCTCAACTGCAACAACCCCTACTATGCCCAATTCAGCGGGAGCAGTTAGAGCGGT 240
|||
Db 36530 GAAATCTCAACTGCAACAACCCCTACTATGCCCAATTCAGCGGGAGCAGTTAGAGCGGT 36471
|||
QY 241 CATCAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 300
|||
Db 36470 CATGAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 36411
|||
QY 301 AGGACTAGCTGATTTCTTAGGCCCAACGAAGAATCCCTAAGCTAGCTGGGAAGGTGACT 360
|||
Db 36410 AGGACTAGCTGATTTCTTAGGCCCAACGAAGAATCCCTAAGCTAGCTGGGAAGGTGACT 36351
|||
QY 361 GCATCCACCTCTAAACATGCGGCTTGCAACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
|||
Db 36350 GCATCCACCTCTAAACATGCGGCTTGCAACTTAGCTCAACCCGACCAATCAGAGAGCTC 36291
|||
QY 421 ACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
|||
Db 36290 ACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCGCTG 36231
|||
QY 481 AGAGCACAGCGGAGGACCAAGGATCGGATATAAACCCAGGCATTTCAGAGCGGCAACGG 540
|||
Db 36230 AGAGCACAGCGGAGGACCAAGGATCGGATATAAACCCAGGCATTTCAGAGCGGCAACGG 36171
|||
QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
|||
Db 36170 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 36111
|||
QY 601 ATTAATCTTGAACCTGAA 619
|||
Db 36110 ACTAAATCTTGAACCTGCA 36092
|||
```

```
RESULT 16
AP001674/c
LOCUS
DEFINITION
Homo sapiens genomic DNA, chromosome 21q, section 18/105.
ACCESSION
AP001674 AL163219 BA000005
VERSION
AP001674.1 GI:7768666
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Rump,A., Schillhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyma,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstieck,G.,
Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichert,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Rieselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
```

Gardiner, K., Nizetic, D., Francis, F., Leirach, H., Reinhardt, R. and Yaspo, M.L.

The DNA sequence of human chromosome 21
Nature 405 (6784), 311-319 (2000)

20289799

10830953

2 (bases 1 to 340000)

AUTHORS

Hattori, M., Fujiyama, A., Taylor, T.D., Watanabe, H., Yada, T., Park, H.S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D.K., Soeda, E., Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K., Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R., Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., Schudy, A., Zimmermann, W., Rosenthal, A., Kudoh, J., Shimizu, N., Nordsiek, G., Asakawa, S., Shintani, A., Sasaki, T., Nagamine, K., Mitsuyama, S., Antonarakis, S.E., Minoshima, S., Shimizu, N., Schoen, O., Desario, A., Hornischer, K., Barand, P., Scharfe, M., Ramser, J., Beck, A., Klages, S., Reichelt, J., Kauer, G., Bloeker, H., Ramser, J., Beck, A., Klages, S., Hennig, S., Riesselmann, L., Dagand, E., Wehrmaier, S., Borzym, K., Gardiner, K., Nizetic, D., Francis, F., Leirach, H., Reinhardt, R. and Yaspo, M.L.

Direct Submission

Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing Consortium: * RIKEN Genomic Sciences Center, Human Genome Research Group * Institute of Molecular Biotechnology, Genome Analysis * Keio University School of Medicine, Dept. of Molecular Biology * GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular Genetics (addresses see below)

On May 30, 2000 this sequence version replaced gi:7717271.

The chromosome 21 mapping and sequencing consortium consisting of

* RIKEN Genomic Sciences Center, Human Genome Research Group, * Sagamihara 228-8555, Japan,

* e-mail: hattori@gsc.riken.go.jp

* URL: http://hgp.gsc.riken.go.jp/

and

* Institute of Molecular Biotechnology, Genome Analysis, * Beutenbergstrasse 11, D-07745 Jena, Germany,

* e-mail: gscj-submit@genome.imb-jena.de

* URL: http://genome.imb-jena.de/

and

* Keio University School of Medicine, Molecular Biology, * Tokyo

160-8582, Japan,

* e-mail: nehimizu@dm-b-med.keio.ac.jp

* URL: http://www.dmb.med.keio.ac.jp/

and

* GBF, Dept. of Genome Analysis,

* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e-mail:

info.genome@gbf.de

* URL: http://genome.gbf.de/

and

* Max-Planck Institute for Molecular Genetics,

* Innestrasse 73, D-14195 Berlin, Germany,

* e-mail: info-chr21@molgen.mpg.de

* URL: http://chr21.rz-berlin.mpg.de/

AL163219: Submitted (10-Apr-2000).

Location/Qualifiers

1. 340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

<1. .6149

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R749P18, 5' partial"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP001537"

<1. .165231

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R853K11, 5' partial"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP001538"

136796. .315056

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R845I12"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AF245226"

156433. 277809

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="P646P19"

/clone_lib="RPCI1.3-5 PAC library"

/note="Accession No. AL157359"

207594. .250588

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clones="Q3D19"

/clone_lib="LL21NCO2-Q Cosmid library"

/note="Accession No. AL049753"

276958. .301481

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clones="R755I19"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP001345"

299634. .>340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R28M9, 3' partial"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP000431"

311751. .>340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="C103E2331, 3' partial"

/clone_lib="ICRFcl03 Cosmid library"

/note="Accession No. AL132993"

15. .444

/note="L1MA4A"

/rpt_family="LINE/L1"

/rpt_type="DISPERSED"

555. .3531

/note="L1MA4A"

/rpt_family="LINE/L1"

/rpt_type="DISPERSED"

3770. .3876

/note="L2"

/rpt_family="LINE/L2"

/rpt_type="DISPERSED"

complement(5645. .6178)

/note="L1MA2"

FEATURES

source

source

source

repeat_region

repeat_region

repeat_region

repeat_region

```
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
6181..6534
/notes="THE18"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
8021..8370
/notes="THE18"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(8373..10032)
/notes="LiMA2"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
10213..10520
/notes="AluDb"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
15810..15851
/notes="(CA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(16628..16697)
/notes="MER53"
/rpt_family="DNA"
/rpt_type=DISPERSED
16638..16715
/notes="MER53"
/rpt_family="DNA"
/rpt_type=DISPERSED
18847..19585
/notes="L2"
/rpt_family="LINE/L2"

Query Match      87.7%; Score 556.6; DB 9; Length 340000;
Best Local Similarity 93.7%; Pred. No. 2.3e-164;
Matches 580; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 CCCTGATCTTTAACTCTTGTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
DB 27921 CCCTGATCTTTAACTCTTGTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 27862
QY 61 CAAATGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 27861 CAAATGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 27802
QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 27801 CTAGACTGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 27742
QY 181 GAAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGAGACAGTGTAGAGCGGT 240
DB 27741 GAAATCAACTGCCCCAACCCCTACTATGCCCAATTCAGCGGAGACAGTGTAGAGCGGT 27682
QY 241 CATCAGCCACCTCCCAACAGCAGTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 300
DB 27681 CATGGCCACCTCCCAACAGCAGTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 27622
QY 301 AGGACTAGCTGGATTTCTTAGGCAACAGGATCCCTAAGCCCTAGCTGGGAGGAGTGA 360
DB 27621 AGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCTAAGCCCTAGCTGGGAGGAGTGA 27562
QY 361 GCATCGACCTCTAAACATAGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGCTC 420
DB 27561 GCATCGACCTCTAAACATAGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGCTC 27502
QY 421 ACTAAATGCTAATTAGGCAAAATAGAGGTAAAGAAATAGCAATCATCTATTGCTG 480
DB 27501 ACTAAATGCTAATTAGGCAAAATAGAGGTAAAGAAATAGCAATCATCTATTGCTG 27442
QY 481 AGAGCAGCGGGGAGGACAGGATCGGGATATAAACCAGGATTCGAGCCCGGACAGG 540
DB 27441 AGAGCAGCGGGGAGGACAGGATCGGGATATAAACCAGGATTCGAGCCCGGACAGG 27382

QY 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATTCTACTCT 600
DB 27381 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATTCTACTCT 27322
QY 601 ATTAAATCTTGCACACTGAA 619
DB 27321 ACTAAATCTTGCACACTGCA 27303

RESULT 17
AP001545 149755 bp DNA linear HTG 13-JUL-2000
Homo sapiens chromosome 18 clone RP11-762G24 map 18q21, WORKING
DRAFT SEQUENCE, 13 unordered pieces.
AP001545
AP001545.3 GI:9189503
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 149755)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
Homo sapiens 149,755 genomic DNA of 18q21
Published Only in Database (2000)
2 (bases 1 to 149755)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
Direct Submission
Submitted (24-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Jul 14, 2000 this sequence version replaced gi:8117383.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: RP11-762G24
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 145596 bases at least Q40
Consensus quality: 147102 bases at least Q30
Consensus quality: 147971 bases at least Q20
Insert size: 148555; sum-of-contigs
Quality coverage: 10.98x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
13 contigs. The true order of the pieces is not known and the
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 44828 contig of 44828 bp in length
44929 82602 contig of 37674 bp in length
82703 101344 contig of 18642 bp in length
101445 116205 contig of 14761 bp in length
116306 123784 contig of 7479 bp in length
123885 130355 contig of 6471 bp in length
130456 133619 contig of 3164 bp in length
133720 137416 contig of 3697 bp in length
137517 141540 contig of 4024 bp in length
141641 144351 contig of 2711 bp in length
144452 146667 contig of 2216 bp in length
```

146768 148219 contig of 1452 bp in length
148320 149755 contig of 1436 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

```

misc_feature      1. .44828
                  /note="assembly_fragment"
misc_feature      44929. .82602
                  /note="assembly_fragment"
misc_feature      82703. .101344
                  /note="assembly_fragment"
misc_feature      101445. .116205
                  /note="assembly_fragment"
misc_feature      116306. .123784
                  /note="assembly_fragment"
misc_feature      123885. .130355
                  /note="assembly_fragment"
misc_feature      130456. .133619
                  /note="assembly_fragment"
misc_feature      133720. .137416
                  /note="assembly_fragment"
misc_feature      137517. .141540
                  /note="assembly_fragment"
misc_feature      141641. .144351
                  /note="assembly_fragment"
misc_feature      144452. .146667
                  /note="assembly_fragment"
misc_feature      146768. .148219
                  /note="assembly_fragment"
misc_feature      148320. .149755
                  /note="assembly_fragment"

```

Query Match 86.9%; Score 551.8; DB 2; Length 149755;
Best local similarity 91.8%; Pred. No. 6.8e-163;

	Matches	583;	Conservative	0;	Mismatches	52;	Indels	0;	Gaps	0;
QY	1	CCCTGTATCTTTAA	ACCTCCTTGTAA	GTTCCTCTCTCCAGAA	TCAAAACTGTAAAA	ACTA	60			
Db	67067	CCCTGTATCTTTAA	ACCTCCTTGTAA	GTTCCTCTCTCCAGAA	TCAAAACTGTAAAA	ACTA	67126			
QY	61	CAAAATTGTTCTT	CAAAATGGAG	ACACAGATGGAG	TCCATGACTAAGAT	TCCA	120			
Db	67127	CAAAATCATTTCT	CAAAATGGAG	CCCCAGATGCAGT	TCCATGACTAAGAT	TCTACCA	67186			
QY	121	CTGGACCGGCTGT	CTAGCCCATGCT	CCGATGTTTAAT	GACATTTGAAG	CACCCCTCCCGAG	180			
Db	67187	CTGGACCAAGCTG	CTAGCCCATGCT	CCGATGTTTAAT	GACATTCGAAG	CAACCCCTCCCAAG	67246			
QY	181	GAAATCTCAATG	CAACACCCCTACT	ATGCCCCCAATTCAG	CGGGAACGACTT	AGAGCGGT	240			
Db	67247	GAAATCTCAACT	GCACAACCCCTACT	ATGCCCCCAAGTT	CAGCAGG	AGCAGTTAGGCGGT	67306			
QY	241	CATCAGCCAACT	CTCCCCAACAG	CACTTGGGTTTTCT	GTGAGAGGGGG	GACTGAGAGAC	300			
Db	67307	CATCAGCCAACT	CTCCCCAACAGCA	CTTGGGTTTTCT	GTGTTCAGAGGGGG	GACTGAGAGAC	67366			
QY	301	AGGACTAGCTGAT	TCTTAGGCCAA	CGAAGAACTCCCT	TAAGCCTTAGCT	TGGGAAGTCACT	360			
Db	67367	AGGACTAGCTGAT	TCTTAGGCTGACT	TAAGAACTCCCT	TAAACCTTAGCT	TGGGAAGTCACT	67426			
QY	361	GCATCCACCTCT	TAAACATGGG	CGTTGCAACTTAG	CTCAACCCGACCA	ATCAGAGACTC	420			
Db	67427	GCATTCACCTTT	TAAACAGAGG	CTTGCAACTTAG	CTCAACCTCAGAC	AGACTC	67486			
QY	421	ACTAAAATGCTA	TTATAGGCAAAAT	TAGAGGTAA	GAANAATGCCNA	TATCTATTGGCTG	480			
Db	67487	ACTAAAATGCTA	TTATAGGCAAA	AAACAGAGG	GTAAAGAAATAG	CCAAATCATCTATTGGCTG	67546			
QY	481	AGAGCACAGGGG	AGGACAGAGT	TCGGGATATAA	CCAGGCAATTCAG	AGCCGCAACGG	540			
Db	67547	AGAGCACAGTGG	GAGGGAACAGAT	CAGGATATAA	CCACAGGCA	TTCAGAGCCGCAACGG	67606			
QY	541	CAACCCCTCTT	GGGTCCCTCCCT	TGTGTATGGG	CGCGCTCTGTTT	TTCACCTCTATTTCAC	600			
Db	67607	CAACCCCTCTT	GGGTCCCTCCCT	TGTGTATGGG	AGCTCTGTTT	TTCACCTCTATTAAATCTT	67666			
QY	601	ATTTAAATCTTG	CAACTGAAAA	AAAAA	AAAAA	AAAAA	635			
Db	67667	TCAACTGCAAAA	ATAATAAATAA	AAAAATATAAAA	ATATAAAA	67701				

RESULT	18
AC090313/c	
LOCUS	152980 bp DNA linear HTG 27-MAR-2003
DEFINITION	Homo sapiens chromosome 18 clone RP11-762G24 map 18.
ACCESSION	AC090313
VERSION	AC090313.12 GI:221233575
KEYWORDS	HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 152980) Birren,B., Nusbaum,C. and Lander,E. Homo sapiens chromosome 18, clone RP11-762G24 Unpublished 2 (bases 1 to 152980) Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	

ORIGIN

McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, J., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
 Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
 Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
 Sougneux, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 152980)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgaiter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, V.,
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 6, 2002 this sequence version replaced gi:21699527.
 All repeats were identified using RepeatMasker:
 Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L12582
 Center clone name: 762_G_24

- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 1 contigs. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the submitter.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.
- * 1 152980: contig of 152980 bp in length.

FEATURES

source
 1. 152980
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"
 /clone="RP11-762G24"
 /clone_lib="RP11 Human Male BAC"

ORIGIN

Query Match 86.9%; Score 551.8; DB 2; Length 152980;
 Best Local Similarity 91.8%; Pred. No. 6.8e-163;
 Matches 583; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTTAACTCCTCTTGTGTTTGTCTTCTCAGAAATCAAAACCTGTAAACTA 60
 |||||
 Db 48688 CCCTGTATCTTTTAACTCCTCTTGTGTTTGTCTTCTCAGAAATCAAAAGCTGTAAACTA 48629
 |||||
 QY 61 CAAATGTTCTTCAATGGAGACAGATGGAGTGCATGACTTAAGTCCACCGTGGACCC 120
 |||||
 Db 48628 CAAATCATTTTCAATGGAGACAGATGGAGTGCATGACTTAAGTCCACAGACCC 48569
 |||||
 QY 121 CTGGAGCGGCTGTAGCCATGCTCCGATGTTTAATGACATTTGAAGGACCCCTCCCGAG 180
 |||||
 Db 48569 CTGGAGCGGCTGTAGCCATGCTCCGATGTTTAATGACATTTGAAGGACCCCTCCCGAG 48509
 |||||
 QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGAGCTAGAGCGGT 240
 |||||
 Db 48508 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGAGCTAGAGCGGT 48449
 |||||
 QY 241 CATCAGCCAACTCCCAACAGCACTTGGTTTCTGTTTGAAGGGGGGAGCTGAGAGAC 300
 |||||
 Db 48448 CATCAGCCAACTCCCAACAGCACTTGGTTTCTGTTTGAAGGGGGGAGCTGAGAGAC 48389
 |||||
 QY 301 AGGACTAGCTGGATTTCTTAGGCCCAAGAAATCCCTAAGCTTAGCTGGGAGGTGACT 360
 |||||
 Db 48388 AGGACTAGCTGGATTTCTTAGGCCCTGACTAAGAAATCCCTAAGCTTAGCTGGGAGGTGACT 48329
 |||||
 QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCGCAATTCAGAGAGCTC 420
 |||||
 Db 48328 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCGCAATTCAGAGAGCTC 48269
 |||||
 QY 421 ACTAAATGCTTAATTAGGCAAAATAGAGGTAAAGAAATGCAATCATCTATTTCCTG 480
 |||||
 Db 48268 ACTAAATGCTTAATTAGGCAAAATAGAGGTAAAGAAATGCAATCATCTATTTCCTG 48209
 |||||
 QY 481 AGAGCACACGGGAGGACAGGATCGGGATATAAACCCAGGCAATTCAGCGCGGCAACGG 540
 |||||
 Db 48208 AGAGCACACGGGAGGACAGGATCGGGATATAAACCCAGGCAATTCAGCGCGGCAACGG 48149
 |||||
 QY 541 CAACCCCTTTTGGGCTCCCTCTCTTGTATGGCGCTCTGTTTCACTCTATTTCCTCT 600
 |||||
 Db 48148 CAACCCCTTTTGGGCTCCCTCTCTTGTATGGCGCTCTGTTTCACTCTATTTCCTCT 48089
 |||||
 QY 601 ATTAATCTTGCACATGAAAAAATAAAAAATAAAAAATAAAAA 635
 |||||
 Db 48088 TCAACTGCAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 48054
 |||||

RESULT 19

AC093531/c 163803 bp DNA linear PRI 16-NOV-2001
 LOCUS Homo sapiens chromosome 5 clone RP11-405L7, complete sequence.
 DEFINITION AC093531
 ACCESSION AC093531
 VERSION AC093531.2 GI:16945981
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 163803)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 163803)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-2001) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 163803)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 COMMENT On Nov 16, 2001 this sequence version replaced gi:15383820.
 Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.3.
Location/Qualifiers
1. .163803
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-405L7"

FEATURES
source

ORIGIN
Query Match 86.9%; Score 551.8; DB 9; Length 163803;
Best Local Similarity 93.2%; Pred. No. 6.9e-163;
Matches 577; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
Db CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATGAAAGCTGTAAAGCTA 117289

QY 61 CAATTTGTTCTTCAATGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db CAATAGTATCTTCAATGAGAACCCAGATGCGAGTCCATGACTAATATCTACCGTGGACCC 117229

QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCCACCCCTCCCGAG 180
Db CTGGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTTGAAGTCAACCCCTCCCGAG 117169

QY 181 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
Db GAAATCTCAACTGACCAACCCCTACTATGCCCAATTCAGTATGAGGAGCGAGTTAGAGCAGT 117109

QY 241 CATCAGCCCACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGAATGAGAGAC 300
Db TGTGAGCCCACTCCCAACAGCAGTACTTGGGTTTCTGTTGAGAGGGTGGATGAGAGAC 117049

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAAGCTTAGCTGGGAGGTGACT 360
Db AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCTAAGCTTAGCTGGGAGGTGACC 116989

QY 361 GCATCCACTCTTAAACATGGGCTTCAACTTAGCTCAGCCGACCAATCAGAGAGCTC 420
Db GCATCCACTCTTAAACATGGGCTTCAACTTAGCTCAGCCGACCAATCAGAGAGCTC 116929

QY 421 ACTAAATGCTAATTAGGCAGAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCGCTG 480
Db ACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCCAAATCATCTATTGCGCTG 116869

QY 481 AGAGCAGCGGAGGAGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGGAACGG 540
Db AGAGCAGCGGAGGAGCAAGGATCGGGATATAAACCCAGGCAATTCAGCCAGCAACAG 116809

QY 541 CAACCCCTTTGGTCCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db CAACCCCTTTGGTCCCTCCCTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 116749

QY 601 ATTAATCTTGCACACTGAA 619
Db ATTAATCTTGCACACTGCA 116730

RESULT 20
AC107075/c
LOCUS AC107075 176425 bp DNA linear PRI 12-JUN-2002
DEFINITION Homo sapiens BAC clone RP11-452N17 from 2, complete sequence.
ACCESSION AC107075
VERSION AC107075.4 GI:21263352
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 176425)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 176425)
Swearengen-Shahid,S., Meyer,R. and Dignan,G.
The sequence of Homo sapiens BAC clone RP11-452N17
Unpublished (2001)
3 (bases 1 to 176425)
Waterston,R.H.
Direct Submission
Submitted (15-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 176425)
Waterston,R.H.
Direct Submission
Submitted (30-MAY-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 176425)
Waterston,R.
Direct Submission
Submitted (12-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 30, 2002 this sequence version replaced gi:18640713.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_NH0452N17

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-624P4; the clone sequenced
to the right is RP11-362J3. Actual start of this clone is at base
position 1 of RP11-452N17; actual end is at base position 176425 of
RP11-452N17.

Polymorphisms exist between AC110086 and AC107075. Data from AC110086, AC116626 and AC018872 was used to finish AC107075.

FEATURES

Source

1. 176425

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-452N17"

/clone_lib="RPCI-11"

2423. 2523

/rpt_family="MIR"

3564. 3598

/rpt_family="AT_rich"

4522. 4567

/rpt_family="AT_rich"

4679. 4756

/rpt_family="(TATAA)n"

5700. 5733

/rpt_family="(A)n"

5973. 6006

/rpt_family="(A)n"

6123. 6144

/rpt_family="AT_rich"

6803. 6896

/rpt_family="L1"

6923. 6999

/rpt_family="L1"

7230. 7715

/rpt_family="MaLR"

7852. 7966

/rpt_family="Tc2"

7966. 8026

/rpt_family="Tc2"

8043. 8172

/rpt_family="Tc2"

8174. 8247

/rpt_family="(TATATG)n"

8250. 8456

/rpt_family="Tc2"

8922. 9642

/rpt_family="L1"

11590. 11632

/rpt_family="AT_rich"

12173. 12321

/rpt_family="MER1_type"

12323. 12343

/rpt_family="(A)n"

13107. 14677

/rpt_family="L1"

14678. 14704

/rpt_family="(TTTG)n"

14708. 14996

/rpt_family="Alu"

14997. 15173

/rpt_family="L1"

15615. 15720

/rpt_family="CT-rich"

15811. 16031

/rpt_family="MaLR"

16093. 16330

/rpt_family="L1"

16724. 17718

/rpt_family="L1"

18878. 19038

/rpt_family="Alu"

19298. 19551

/rpt_family="Alu"

19552. 19912

/rpt_family="MaLR"

23674. 23705

/rpt_family="(CAGA)n"

repeat_region 25248. 25839
/rpt_family="MaLR"
repeat_region 27499. 27532
/rpt_family="AT_rich"
repeat_region 27693. 27840
/rpt_family="MER1_type"
repeat_region 27902. 28090
/rpt_family="MIR"
repeat_region 28091. 28254
/rpt_family="MER1_type"
repeat_region 28283. 28477
/rpt_family="MER1_type"
repeat_region 28510. 28533
/rpt_family="AT_rich"
repeat_region 28629. 28649
/rpt_family="AT_rich"
repeat_region 29707. 29751
/rpt_family="AT_rich"
repeat_region 30414. 30847
/rpt_family="(TA)n"
repeat_region 31385. 31942
/rpt_family="L1"
repeat_region 32421. 32591
/rpt_family="L2"
repeat_region 32592. 32983
/rpt_family="MaLR"
repeat_region 32984. 33714
/rpt_family="L2"
repeat_region 34913. 35339
/rpt_family="MaLR"
repeat_region 35752. 36049
/rpt_family="Alu"
repeat_region 36274. 36330
/rpt_family="Alu"
repeat_region 36767. 36809
/rpt_family="AT_rich"
repeat_region 37396. 37437
/rpt_family="AT_rich"
repeat_region 37816. 37850
/rpt_family="A-rich"
repeat_region 38954. 39375
/rpt_family="L1"
repeat_region 40130. 40216
/rpt_family="L2"
repeat_region 41723. 42015
/rpt_family="Alu"

Query Match 86.9%; Score 551.8; DB 9; Length 176425;

Best Local Similarity 92.8%; Pred. NO. 6.9e-163;

Matches 593; Conservative 0; Mismatches 37; Indels 9; Gaps 1;

QY	1	CCCTGTATCTTTAACTCCTCTGTTAGTTTGTCTTCCAGAACTCAAACTGTAACCTA	60
Db	161068	CCCTATATCTTTAACTCCTCTGTTAGTTTGTCTTCCAGAACTCAAACTGTAACCTA	161009
QY	61	CAAAATGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC	120
Db	161008	CAAAATGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC	160949
QY	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG	180
Db	160948	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG	160889
QY	181	GAAATCTCAACTGCACAACCCCTACTATGCCCAATTCAGCGGACGAGTGTAGACGGT	240
Db	160888	GAAATCTCAACTGCACAACCCCTACTATGCCCAATTCAGCGGACGAGTGTAGACGGT	160829
QY	241	CATCAGCAACCTCCCAACAGACATTTGGGTTTTCTCTTTGAGAGGGGAGCTGAGAGAC	300
Db	160828	CTTCGCCCAACCTCCCAACAGACATTTGGGTTTTCTCTTTGAGAGGGGAGCTGAGAGAC	160769
QY	301	AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAGCCCTAGCTGGGAAGGTGACT	360

Fri Feb 25 16:26:29 2005

us-09-319-156b-6.rge

Db 160768 AGGACTAGCTGGATTTCCTAGCGCGACTAAGATCCCTAACGCTAGCTGGGAAGTGACC 160709
Qy 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCTACACCGACCAATCA----- 412
Db 160708 GCTTCATCTTTAAACACGGGGCTTACAACTTAACCTCACATGACCAATCAGATAGTAA 160649
Qy 413 -GAGAGCTCACTAAATGCTAATTAGGCAGAAAATAGGAGGTAAAGAAATACCCATCATC 471
Db 160648 GGAGAGCTCACTAAATGCTAATTAGGCAGAAAATAGGAGGTAAAGAAATACCCATCACC 160589
Qy 472 TATTCCTGAGACAGCAGCGGAGGAGGACAGGATCGGATATAAACCCAGGCAATTCGAGC 531
Db 160529 TGTTCCTGAGACAGCAGCGGAGGAGGACATGATCGGATATAAACCCAGGCAATTCGAGC 160529
Qy 532 CGGCAACGGCAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTA 591
Db 160528 CGGCAACGACAAACCCCTTTGGGTCCTCCCTCTTGTATGGGAGCTGCTTTTCACTCTA 160469
Qy 592 TTTCACTCTAATTAATCTTGCAACTGAAAAAATAAAAA 630
Db 160468 CTTCACTCTATCAATCTTTGCAACTGCAAAAAAATAAAAA 160430

RESULT 21
AL139090/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-12B13 on chromosome 6 Contains
STSs and GSSs, complete sequence.
ACCESSION
AL139090
VERSION
AL139090.11 GI:11228536
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 128468)
Garner, P.
Direct Submission
Submitted (05-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 20, 2000 this sequence version replaced gi:11125403.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Swi., SWISSPROT; Tr., TREMBL; Wp., WormPeP; Information
on the WormPeP database can be found at
http://www.sanger.ac.uk/projects/C-elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
IMPORTANT: This sequence is not the entire insert of clone
RP11-12B13 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-12B13 is at 1 in this sequence. The
true left end of clone RP1-125N20 is at 128369 in this sequence.
The true right end of clone RP11-13D18 is at 88802 in this
sequence. This sequence was finished as follows unless otherwise
noted: all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
RP11-12B13 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm

FEATURES
source
VECTOR: pBACE3.6.
Location/Qualifiers
1..128468
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-12B13"
/clone_lib="RPCI-11.1"
1..353
/notes="match: GSS: Em:AQ609536"
7..448
/notes="match: GSS: Em:AQ704228"
68..350
/notes="match: GSS: Em:B75711"
571..1535
/notes="LRRS repeat: matches 1..966 of consensus"
1966..2087
/notes="L2 repeat: matches 2574..2704 of consensus"
2290..2509
/notes="L2 repeat: matches 2098..2328 of consensus"
2976..3352
/notes="L2 repeat: matches 1693..2089 of consensus"
3497..3688
/notes="MER20 repeat: matches 16..188 of consensus"
complement(4521..5085)
/notes="match: GSS: Em:AQ421033"
complement(4571..5086)
/notes="match: GSS: Em:AQ883615"
complement(4633..5083)
/notes="match: GSS: Em:AQ812769"
5012..5530
/notes="MER41B repeat: matches 114..601 of consensus"
5081..5634
/notes="match: GSS: Em:AQ544837"
6085..6174
/notes="MIR repeat: matches 64..153 of consensus"
6235..6416
/notes="MER5A repeat: matches 3..186 of consensus"
7146..8805
/notes="L1ME3 repeat: matches 4381..6140 of consensus"
8806..9399
/notes="L1PA5 repeat: matches 5526..6143 of consensus"
9400..9710
/notes="L1ME3 repeat: matches 4094..4381 of consensus"
9711..15846
/notes="L1PA6 repeat: matches 11..6143 of consensus"
15826..16269
/notes="L1M4 repeat: matches 3692..4135 of consensus"
17077..17544
/notes="L1M4 repeat: matches 2710..3168 of consensus"
17552..17740
/notes="L1MEC repeat: matches 2408..2260 of consensus"
17741..18039
/notes="AluX repeat: matches 1..299 of consensus"
18040..18346
/notes="L1MEC repeat: matches 2106..2409 of consensus"
18447..18874
/notes="L1MEC repeat: matches 1469..1910 of consensus"
complement(18475..18995)
/notes="match: GSS: Em:AQ755181"
18938..19072
/notes="L1MEC repeat: matches 1243..1377 of consensus"
19459..19514
/notes="28 copies 2 mer ca 96% conserved"
20321..20675
/notes="LTR16B repeat: matches 97..464 of consensus"
20952..21137
/notes="AluJ repeat: matches 120..307 of consensus"
21203..21254
/notes="26 copies 2 mer aa 71% conserved"
23289..23682
/notes="match: GSS: Em:AQ881909"

repeat_region 23421..23472
 /note="MIR repeat: matches 90..141 of consensus"
 misc_feature 24395..24791
 /note="match: GSS: Em:AQ027280"
 repeat_region 24638..24848
 /note="HAL1 repeat: matches 428..638 of consensus"
 repeat_region 25008..25217
 /note="105 copies 2 mer aa 55% conserved"
 repeat_region 23328..25391
 /note="32 copies 2 mer ta 70% conserved"
 repeat_region 25482..26261
 /note="LIP3 repeat: matches 5..776 of consensus"
 repeat_region 26257..31504
 /note="LIP3 repeat: matches 900..6146 of consensus"
 repeat_region 31614..32312
 /note="LIM4 repeat: matches 3918..4608 of consensus"
 repeat_region 32398..32650
 /note="LIM4 repeat: matches 4626..4909 of consensus"
 repeat_region 32799..32963
 /note="FRAM repeat: matches -2..162 of consensus"
 repeat_region 33151..33174
 /note="12 copies 2 mer tt 95% conserved"
 repeat_region 33176..33500
 /note="LIR17 repeat: matches 1..326 of consensus"
 repeat_region 33501..33791
 /note="HERV17 repeat: matches 8244..8523 of consensus"
 repeat_region 33811..38137
 /note="HERV17 repeat: matches 1651..5933 of consensus"
 repeat_region 38135..39957
 /note="HERV17 repeat: matches 1..1836 of consensus"
 repeat_region 39958..40489
 /note="LIR17 repeat: matches 257..780 of consensus"
 repeat_region 40490..41321
 /note="LIME3A repeat: matches 5261..6131 of consensus"
 repeat_region 41354..41655
 /note="AluY repeat: matches 1..302 of consensus"
 repeat_region 41971..42279
 /note="AluX repeat: matches 1..303 of consensus"
 repeat_region 43255..43365
 /note="LIM4 repeat: matches 2169..2266 of consensus"
 repeat_region 43366..43774
 /note="MSTA repeat: matches 1..426 of consensus"
 repeat_region 43775..44133
 /note="LIM4 repeat: matches 2266..2691 of consensus"
 repeat_region 44140..44339
 /note="LIM4 repeat: matches 6093..6289 of consensus"
 repeat_region 46709..47216
 /note="LIMB8 repeat: matches 5658..6169 of consensus"
 repeat_region 47217..47431
 /note="LIM4 repeat: matches 4637..4856 of consensus"
 repeat_region 47487..47839
 /note="MLT2E repeat: matches 8..345 of consensus"
 repeat_region 47848..48031
 /note="LIM4 repeat: matches 4451..4645 of consensus"
 repeat_region 48017..49417
 /note="LIMEC repeat: matches 2373..3522 of consensus"
 repeat_region 49436..50712
 /note="LIMD2 repeat: matches 5056..6333 of consensus"
 misc_feature complement(50811..51360)
 /note="match: GSS: Em:AQ536659"
 repeat_region 52289..56182
 /note="LIP3 repeat: matches 2255..6146 of consensus"
 repeat_region 56710..57178
 /note="MLT1C repeat: matches 1..463 of consensus"
 repeat_region 57746..57868
 /note="MER33 repeat: matches 202..323 of consensus"
 repeat_region 57869..58554
 /note="MER44C repeat: matches 1..727 of consensus"
 repeat_region 58555..58752
 /note="MER33 repeat: matches 1..202 of consensus"
 repeat_region 58890..59200
 /note="AluSq repeat: matches 1..307 of consensus"
 repeat_region 59630..60070

misc_feature /note="LIMB3 repeat: matches 5737..6184 of consensus"
 59956..60646
 /note="match: GSS: Em:B82923"
 repeat_region 61550..61582
 /note="LIMB2 repeat: matches 6133..6164 of consensus"
 repeat_region 61583..62103
 Query Match 86.8%; Score 551.4; DB 9; Length 128468;
 Best Local Similarity 93.9%; Pred. No. 8.8e-163;
 Matches 596; Conservative 0; Mismatches 36; Indels 3; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGATCAAACTCAAACTCAAACTCA 60
 Db CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGATCAAACTCAAACTCAAACTCA 33732
 QY 61 CAAATGTTCTTCAATGAGCACCAGATGGATGATGATGATGATGATGATGATGATGATG 120
 Db CAAATGTTCTTCAATGAGCACCAGATGGATGATGATGATGATGATGATGATGATGATG 33732
 QY 121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAAATGACATTAAGGACACCCCTCCCGAG 180
 Db CT-GATCAACCTCTTAGCCCATGCTCCGATGTTAAATGACATTAAGGACACCCCTCCCGAG 33613
 QY 181 GAAATCTCAACTGCAACCCCTTACTATGCCCAATTTAGCGGGAGGAGCTTAGACCGGT 240
 Db GAAATCTCAACTGCAACCCCTTACTATGCCCAATTTAGCGGGAGGAGCTTAGACCGGT 33553
 QY 241 CATCAGCNAACCTCCCAACAGCACTTGGTTTCTGTTGAGAGGGGAGCTTGAGAGAC 300
 Db CATCAGCNAACCTCCCAACAGCACTTGGTTTCTGTTGAGAGGGGAGCTTGAGAGAC 33493
 QY 301 AGGACTAGCTGGATTTCTTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAGGTGACT 360
 Db AGGACTAGCTGGATTTCTTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAGGTGACT 33433
 QY 361 GCATCCACCTTAAACATGGGGCTTGCACCTTAGCTTACACCCGACCAATCAGAGCTC 420
 Db GCATCCACCTTAAACATGGGGCTTGCACCTTAGCTTACACCCGACCAATCAGAGCTC 33373
 QY 421 ACTAAATGCTAATTAGGCANAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
 Db ACTAAATGCTAATTAGGCANAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 33313
 QY 481 AGACACACGGGAGGACAAGGATCGGATATAAACCAGGATTCGAGCCGGCAACGG 540
 Db AGACACACGGGAGGACAAGGATCGGATATAAACCAGGATTCGAGACGCAATGG 33253
 QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCCTCT 600
 Db CAATCCCTTTGGTCCCTCCCTCCCTTTGATGGAGCTCTGTTTCACTCTATTTCCTCT 33195
 QY 601 ATTAATCTTCAACTGAAAAAAGAAAAA 635
 Db ATTAATCTTCAACTGAAAAAAGAAAAA 33160

RESULT 22
 AL139038
 LOCUS Human DNA sequence from clone RP11-456B18 on chromosome 13,
 DEFINITION complete sequence.
 ACCESSION AL139038
 VERSION AL139038.18 GI:14800148
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 140756)
 DUNN,M.
 Direct Submission
 Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone


```

repeat_region 32449..32821
/note="THE1C repeat: matches 1..371 of consensus"
repeat_region 32824..34484
/note="THE1C-internal repeat: matches 3..1651 of
consensus"
repeat_region 34465..34825
/note="THE1C repeat: matches 1..360 of consensus"
repeat_region 34804..35004
/note="MER67C repeat: matches 202..403 of consensus"
repeat_region 3792..39868
/note="LIPAL12 repeat: matches 6076..6152 of consensus"
repeat_region 40953..41258
/note="AluY repeat: matches 1..308 of consensus"
repeat_region 41279..41308
/note="15 copies 2 mer aa 86% conserved"
repeat_region 42744..42933
/note="LIMAB repeat: matches 6064..6256 of consensus"
repeat_region 42973..43275
/note="AluSX repeat: matches 1..303 of consensus"
misc_feature complement(43560..43991)
/note="match: SRS: Em:HSPH07E5"
repeat_region 45196..45584
/note="MLT2B repeat: matches 2..399 of consensus"
repeat_region 45594..45767
/note="87 copies 2 mer ta 77% conserved"
repeat_region 45800..45851
/note="MLT2B repeat: matches 394..448 of consensus"
repeat_region 48118..48487
/note="THE1C repeat: matches 1..370 of consensus"
repeat_region 48708..48806
/note="AluJo/FRAM repeat: matches 197..297 of consensus"
repeat_region 49853..50271
/note="LIMEC repeat: matches 2274..2369 of consensus"
repeat_region 50637..50927
/note="LIME repeat: matches 974..1253 of consensus"
repeat_region 52250..52349
/note="MIR repeat: matches 131..228 of consensus"
repeat_region 52696..52996
/note="AluSg repeat: matches 1..293 of consensus"

Query Match 86.8%; Score 551.4; DB 9; Length 140756;
Best Local Similarity 94.3%; Pred. No. 8.9e-163;
Matches 584; Conservative 0; Mismatches 31; Indels 4; Gaps 1;

QY 1 CCTGTATCTTAACTCTTGTAGTCTTCCAGATCAAACTGTAACCTA 60
DB 69181 CCTGTATCTTAACTCTTGTAGTCTTCCAGATCAAACTGTAACCTA 69240
QY 61 CAAATTGCTTCAAAATGGAGCACCAGATGGATGCTTCCAGATCAAACTGTAACCTA 120
DB 69241 CAAATTGCTTCAAAATGGAGCACCAGATGGATGCTTCCAGATCAAACTGTAACCTA 69300
QY 121 CTGACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCGAG 180
DB 69301 CTGACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCGAG 69360
QY 181 GAATCTCACTGACAAACCCCTACTATGCTCCCAATTCAGGGGAGCAGTGTAGCGGT 240
DB 69361 GAATCTCACTGACAAACCCCTACTATGCTCCCAATTCAGGGGAGCAGTGTAGCGGT 69416
QY 241 CATCAGCAACCTCCCAACAGCACTTGGTGTTCCTGTTGAGAGGGGAGCAGTGTAGAG 300
DB 69417 CATGCGCACTTCCCAACAGCACTTGGTGTTCCTGTTGAGAGGGGAGCAGTGTAGAG 69476
QY 301 AGGACTAGTGGATTTCTAGGCCAACGAAGAATCCCTAAGCCCTAGTGTGGGAGGTGACT 360
DB 69477 AGGACTAGTGGATTTCTAGGCCAACGAAGAATCCCTAAGCCCTAGTGTGGGAGGTGACC 69536
QY 361 GCATCCACCTTAAACATGGGGCTTGCACTTACTACACCCCAACCAATCAGAGCTC 420
DB 69537 GCATCCACCTTAAACATGGGGCTTGCACTTACTACACCCCAACCAATCAGAGCTC 69596
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCTG 480

```

```

Db 69597 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCTG 69656
QY 481 AGAGCACACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCAGCGCGGCAACGG 540
DB 69657 AAAGCACACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCAGCGCGGCAACAG 69716
QY 541 CAACCCCTTTGGTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 69717 CAACCCCTTTGGTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 69776
QY 601 ATTAATCTTGAACCTGAA 619
DB 69777 GTTAAATCTTGAACCTGCA 69795

RESULT 23
AC021774/LOCUS 167366 bp DNA linear HTG 30-MAR-2000
DEFINITION Homo sapiens clone RP11-13C18, WORKING DRAFT SEQUENCE, 10 unordered
pieces.
AC021774 GI:7341907
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 167366)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone RP11-13C18
Unpublished
2 (bases 1 to 167366)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Bédar, F.,
Boguslavsky, L., Boukhgaiter, B., Brown, A., Burkett, G., Castle, A.,
Chospel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeArelano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lander, E., Lehotsky, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6940059.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3315
Center clone name: 13 C.18
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161819 bases at least Q40
Consensus quality: 164760 bases at least Q30
Consensus quality: 165688 bases at least Q20
Insert size: 170000; agarose-fp

```

RESULT 24	AL607153	183499 bp	DNA	linear	HTG 21-OCT-2001
LOCUS	AL607153	Homo sapiens chromosome 13 clone RP13-276D12.			
DEFINITION	AL607153				
ACCESSION	AL607153				
VERSION	AL607153.3	GI:15902188			
KEYWORDS	HTG; HTGS PHASE2; HTGS_CANCELLED.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	Lovell, J.			
AUTHORS	Direct Submission				
TITLE	Submitted (18-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,				
JOURNAL	CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk				
	requests: clonerequest@sanger.ac.uk				
	On Oct 5, 2001 this sequence version replaced gi:15962724.				
COMMENT	-----	Genome Center			
	Center: Sanger Centre				
	Center code: SC				
	Web site: http://www.sanger.ac.uk				
	Contact: humquery@sanger.ac.uk				
	-----	Project Information			
	Center project name: hb276D12				
	-----	Summary Statistics			
	Assembly program: XGAP4; version 4.5				
	Sequencing vector: plasmid; L08752; 100% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Consensus quality: 183433 bases at least Q40				
	Consensus quality: 183439 bases at least Q30				
	Consensus quality: 183441 bases at least Q20				
	Insert size: 183499; sum-of-contigs				

	Query Match	86.8%	Score 551.4	DB 2	Length 167366
	Best Local Similarity	93.9%	Pred. No. 9.2e-163		
	Matches 596	Conservative	0	Mismatches 36	Indels 3
					Gaps 2
QY	1	CCCTGTATCTTTAAACCTCCCTTGTGTAAAGTTGTCTCTCCAGAAATCAAAACTGTATAACTA	60		
DB	7535	CCCTGTATCTTTAAACCTCCCTTGTGTAAAGTTGTCTCTCTCCAGAAATGAAACTGTATAACTA	7476		
QY	61	CAAAATTGTTCTTCAAATGGAGCACACAGATGGAGTGCATGACTAGATCCACCGTGGACCC	120		

Insert size: 179011; 5.6% error; agarose-fp
Quality coverage: 12.84x in Q20 bases; sum-of-contigs Quality
coverage: 13.23x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1. 183499: contig of 183499 bp in length.

FEATURES	Location/Qualifiers
source	1. .183499
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="13"
	/clone="RP13-276D12"
	/clone_lib="RP13-2"
misc_feature	1. .183499
	/note="assembly_fragment:05362"
ORIGIN	
Query Match	86.8%; Score 551.4; DB 2; Length 183499;
Best Local Similarity	94.3%; Pred. No. 9.4e-163;
Matches 584; Conservative	0; Mismatches 31; Indels 4; Gaps 1;
QY	1 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCCAGATCAAACTGTAATACTA 60
Db	12044 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCCAGATCAAACTGTAATACTA 12103
QY	61 CAAATTGTTCTTCAATGGAGCCAGATGAGTGCATGACTAAGATCCACCTGGACCC 120
Db	12104 CAAATTGTTCTTCAATGGAGCCAGATGAGTGCATGACTAAGATCCACCTGGACCC 12163
QY	121 CTGACCGGCTGCTAGCCATCTCGATGTTAATGACATTCAGAGCCCTCCCGAG 180
Db	12164 CTGACCGGCTGCTAGCCATCTCGATGTTAATGACATTCAGAGCCCTCCCGAG 12223
QY	181 GAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCAGCGGAGCAGTATGAGCGGT 240
Db	12224 GAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCAGCGGAGCAGTATGAGCGGT 12279
QY	241 CATCAGCAACCTCCCAACAGACACTGGTTTCTGTTGAGAGGGGACTGAGAGAC 300
Db	12280 CATTCGCAACTCTCCCAACAGACACTGGTTTCTGTTGAGAGGGGACTGAGAGAC 12339
QY	301 AGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTAGCTAGCTAGCTAGCT 360
Db	12340 AGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTAGCTAGCTAGCTAGCT 12399
QY	361 GCATCCACCTCTAAACATGGGCTTGCAACTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db	12400 GCATCCACCTCTAAACATGGGCTTGCAACTAGCTCACACCCGACCAATCAGAGAGCTC 12459
QY	421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGGCAATCATCTATTGCTGT 480
Db	12460 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGGCAATCATCTATTGCTGT 12519
QY	481 AGACACAGCGGGAGGACAGGATCGGGATATTAACCCAGGATTCGAGCCGCGACAGG 540
Db	12520 AAAGCAGCGGGAGGACAGGATCGGGATATTAACCCAGGATTCGAGCCGCGACAGG 12579
QY	541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGGCTCTGTGTTTCACTCTATTTCCTCT 600
Db	12580 CAACCCCTTTGGTCCCTCCCTTTGTATGGGGCTCTGTGTTTCACTCTATTTCCTCT 12639
QY	601 ATTAATCTTGCAACTGAA 619
Db	12640 GTTAAATCTTGCAACTGCA 12658

RESULT 25
AP002790/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AP002790 164211 bp DNA linear HTG 13-JUL-2000
Homo sapiens chromosome 18 clone RP11-732P12 map 18q21, WORKING
DRAFT SEQUENCE, 20 unordered pieces.
AP002790
HTG: HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164211)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 164,211 genomic DNA of 18q21
Published Only in Database (2000)
2 (bases 1 to 164211)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (11-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: RP11-732P12
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 154344 bases at least Q40
Consensus quality: 159150 bases at least Q30
Insert size: 162311; sum-of-contigs
Quality coverage: 5.05x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
20 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1
28606 contig of 28606 bp in length
55373 contig of 27267 bp in length
73701 contig of 17628 bp in length
92513 contig of 18712 bp in length
102171 contig of 9558 bp in length
111253 contig of 8982 bp in length
120167 contig of 8814 bp in length
127637 contig of 7370 bp in length
134267 contig of 6530 bp in length
138746 contig of 4379 bp in length
143265 contig of 4419 bp in length
147843 contig of 4278 bp in length
150342 contig of 2599 bp in length
153002 contig of 2560 bp in length
155224 contig of 2122 bp in length
156567 contig of 1243 bp in length
158332 contig of 1665 bp in length
160143 contig of 1711 bp in length
162721 contig of 2478 bp in length
164211 contig of 1390 bp in length
162822

* NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 28606: contig of 28606 bp in length
* 28607: gap of 100 bp
* 28707: contig of 27267 bp in length
* 55973: gap of 100 bp
* 55974: contig of 17628 bp in length
* 56074: gap of 100 bp
* 73701: contig of 18712 bp in length
* 73801: gap of 100 bp
* 73802: contig of 9558 bp in length
* 92513: gap of 100 bp
* 92514: contig of 100 bp
* 92614: gap of 100 bp
* 102171: contig of 8982 bp in length
* 102172: gap of 100 bp
* 102272: contig of 8814 bp in length
* 111253: gap of 100 bp
* 111254: contig of 8814 bp in length
* 120167: gap of 100 bp
* 120168: contig of 7370 bp in length
* 120268: gap of 100 bp
* 127638: contig of 6530 bp in length
* 127738: gap of 100 bp
* 134268: contig of 4379 bp in length
* 134368: gap of 100 bp
* 138747: contig of 4419 bp in length
* 143265: gap of 100 bp
* 143266: contig of 4278 bp in length
* 143366: gap of 100 bp
* 147644: contig of 2599 bp in length
* 147744: gap of 100 bp
* 150343: contig of 2560 bp in length
* 150443: gap of 100 bp
* 153003: contig of 2122 bp in length
* 153102: gap of 100 bp
* 153224: contig of 1243 bp in length
* 155325: gap of 100 bp
* 156567: contig of 1665 bp in length
* 156668: gap of 100 bp
* 158332: contig of 1711 bp in length
* 158433: gap of 100 bp
* 160143: contig of 2478 bp in length
* 160244: gap of 100 bp
* 162722: contig of 1390 bp in length.
* 162822: Location/Qualifiers

FEATURES
source

1. .164211
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q21"
/clone="Rp11-732p12"
1. .28606
/note="assembly_fragment"
28707. .55973
/note="assembly_fragment"
56074. .73701
/note="assembly_fragment"
73802. .92513
/note="assembly_fragment"
92614. .102171
/note="assembly_fragment"
102272. .111253
/note="assembly_fragment"
111354. .120167
/note="assembly_fragment"
120268. .127637
/note="assembly_fragment"
127738. .134267
/note="assembly_fragment"

misc_feature 134368. .138746
/note="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature 138847. .143265
/note="assembly_fragment"
misc_feature 143366. .147643
/note="assembly_fragment"
misc_feature 147744. .150342
/note="assembly_fragment"
misc_feature 150443. .153002
/note="assembly_fragment"
misc_feature 153103. .155224
/note="assembly_fragment"
misc_feature 155325. .156567
/note="assembly_fragment"
misc_feature 156668. .158332
/note="assembly_fragment"
misc_feature 158433. .160143
/note="assembly_fragment"
misc_feature 160244. .162721
/note="assembly_fragment"
misc_feature 162822. .164211
/note="assembly_fragment"

ORIGIN

Query Match 86.6%; Score 550.2; DB 2; Length 164211;
Best Local Similarity 91.7%; Pred. No. 2.2e-162;
Matches 582; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1 CCCTGTATCTTTAACTCCCTTGTAGTTTGTCTCTCCAGATCAAACTGTAATACTA 60
Db 6465 CCCTGTATCTTTAACTCCCTTGTAGTTTGTCTCTCCAGATCAAACTGTAATACTA 6406
QY 61 CAAATTTGTTCTTCAATGGAGCACAGATGAGTCCATGATTAAGATCCACCGTGGACCC 120
Db 6405 CAAATCATCTTCAATGGAGCCCCAGATGAGTCCATGATTAAGATCCACAGACCC 6346
QY 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGCACCCCTCCCGAG 180
Db 6345 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACAGGAGGCAACCTCCCAAG 6286
QY 181 GAAATCTCACTGACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 6285 GAAATCTCACTGACAAACCCCTACTATGCCCAAGTTTCCAGGAGGAGGAGTCTGAGAC 6226
QY 241 CATCAGCAACCTCCCAACAGCATTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
Db 6225 CATCAGCAACCTCCCAACAGCATTTGGGTTTCTGTTGAGAGGGGGTACTGAGAGAC 6166
QY 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAGAATCCCTTAAGCCTAGCTGGGAAGGTGACT 360
Db 6165 AGGACTAGCTGGATTTCTTAGGCTGACTTAAGATCCCTTAACCTAGCTGGGAAGGTGACT 6106
QY 361 GCATCCACTCTAAACATGGGGCTTTGCACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 6105 GCATTCACCTTTAAACACGGGGCTTTGCACTTAGCTCAACCTAGCTGGGAAGGTGACT 6046
QY 421 ACTAAATGCTTAATTAGGCAAAATAGGAGTAAAGATAGCAATATCATCTATTGCGCTG 480
Db 6045 ACTAAATGCTTAATTAGGCAAAACAGGAGTAAAGATAGCAATATCATCTATTGCGCTG 5986
QY 481 AGAGCACAGCGGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCGAGCGGCAACGG 540
Db 5985 AGAGCACAGTGGGAGGAGCAAGGATCAGGATATAAACCAGGCAATTCGAGCGGCAACGG 5926
QY 541 CAACCCCTTTGGGTCCTCCCTCCCTTTGATGGGCGCTCTGTTTTCACCTCTATTTCACCT 600
Db 5925 CAACCCCTTTGGGTCCTCCCTCCCTTTGATGGGAGCTCTGTTTTCACCTCTATTTCACCT 5866
QY 601 ATTAATCTTGCACTGAAAAAATAAAAAA 635
Db 5865 TCAACTGCAAAAAATAAAAAATAAAAAA 5831

RESULT 26

AC064801
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

167843 bp DNA linear PRI 30-JUL-2002
Homo sapiens chromosome 18, clone RP11-120K19, complete sequence.
AC064801
AC064801.10 GI:22004284
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 167843)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-120K19
Unpublished
2 (bases 1 to 167843)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehotsky,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mieng,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (22-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 167843)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouckhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mieng,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tessaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 167843)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Bouckhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mieng,V.,

RESULT	27
CNS01DSH	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

CNS01DSH 180523 bp DNA linear PRI 20-AUG-2001
Human chromosome 14 DNA sequence BAC R-30513 of library RPCI-11
from chromosome 14 of Homo sapiens (human), complete sequence.
AL121784
AL121784.5 GI:15282089
HTG: HTGS-ACTIVERFIN.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 180523)
Heilig,R., Petit,J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brotier, P., Catcollin, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
Guyapay, G., Saurin, W. and Weissenbach, J.
Genotyping of the human chromosome 14
Unpublished
2 (bases 1 to 180523)
Genoscope.
Direct Submission
Submitted (01-AUG-2001) Genoscope - Centre National de Sequençage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Aug 23, 2001 this sequence version replaced gi:12001714.
----- Genome Center
Center: Genoscope / Centre National de Sequençage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : C-2547L24 (AC=AL135810)
Downstream BAC (overlapping the SP6 end) : R-580E16 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.78x in Q20 bases; sum-of-contigs

Overall quality chart :

Range : bases
0 :
1 - 9 :
10 - 19 : 18
20 - 29 : 87
30 - 39 : 360
40 - 49 : 3700
50 - 59 : 10918
60 - 69 : 13753
70 - 79 : 29245
80 - 89 : 58279
90 - 99 : 64163

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES

source

Location/Qualifiers
1. .180523
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone_lib="RP11-11"
1561. .1825
/notes="matching EMBL:R74544
RHdb:RH53560
dbSTS:ST342868
Identified using the e-PCR software (G. Schuler)"
69085. .69312
/notes="matching EMBL:G03465
RHdb:RH53501
RHdb:RH3454
dbSTS:ST31070
Identified using the e-PCR software (G. Schuler)"
71190. .71397
/notes="matching EMBL:AA010373
RHdb:RH75482
dbSTS:ST52570
Identified using the e-PCR software (G. Schuler)"

STS

STS

STS

ORIGIN

Query Match 86.5%; Score 549; DB 9; Length 180523;
Best Local Similarity 92.9%; Pred. No. 5.4e-162;
Matches 590; Conservative 0; Mismatches 35; Indels 10; Gaps 1;
QY 1 CCCTGTATCTTAACCTCTTGTAGTTTGTCTTCCAGATCAAACTGTAAACTA 60
DB 537 CCCTTTATCTTTAACTCTTGTAGTTTGTCTTCCAGATCAAACTGTAAACTA 596
QY 61 CAAATTTCTTCAATGGAGCCACGATGGATGCTTCAATGATCCACCGTGGACCC 120
DB 597 CAAATAGTTCTTCAATGGAGCCACGATGGATGCTTCAATGATCCACCGTGGACCC 656
QY 121 CTGGACCGGCTGTAGCCATGCTCCGATGTTTAAATGACATGAAGCACCCTCCCGAG 180
DB 657 CTGGACCGGCTGTAGCCATGCTCCGATGTTTAAATGACATGAAGCACCCTCCCGAG 716
QY 181 GAAATCTCAATGTCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 717 GAAATCTCAATGTCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 776
QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTTGTGAGAGGGGACTGAGAGAC 300
DB 777 TGTGAGCAACCTCCCAACAGCACTTGGGTTTCTTGTGAGAGGGGACTGAGAGAC 836
QY 301 AGGACTAGCTGGATTTCTTAGGCCAACAGAGATCCCTTAAGCCCTAGCTGGGAAGGTGACT 360
DB 837 AGGACTAGCTGGATTTCTTAGGCCAACAGAGATCCCTTAAGCCCTAGCTGGGAAGGTGACC 896
QY 361 GCATCCACCTCTTAACATGGGGCTTGAACCTAGCTACACCCGACCAATCAGAGAGCTC 420
DB 897 ACATCCACCTTTAAACACAGGGCTTGAACCTAGCTACACCCGACCAATCAGAGAGCTC 956

QY 421 ACTAAATGCTTAATAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
DB 957 ACTAAATGCTTAATAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 1016
QY 481 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGACATTTCAGACCGGCAACGG 540
DB 1017 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGACATTTCAGACCGGCAACGG 1076
QY 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGCGCTCTGTTTTCACCTCTATTTCACCTCT 600
DB 1077 CTACCCCTTTGGGTCCTCCCTCTTGTATGGGAGCTCAGT-----TTTCACTCT 1126
QY 601 ATTAATCTTGAACCTGAAAAAATAAAAAA 635
DB 1127 ATTAATCTTGAACCTGAAAAAATAAAAAA 1161

RESULT 28
AC007374
LOCUS
DEFINITION
Homo sapiens chromosome 14 clone RP11-325L17 map 14q31, complete sequence.
AC007374
VERSION
AC007374.6 GI:13450000
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 190565)
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,
Pate, D. and Hood, L.
Sequencing of human chromosome 15 D15S146-D15S117 region
Unpublished
2 (bases 1 to 190565)
Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
James, R., Loretz, C., Lasky, S., Madan, A., Prescott, S., Ratcliffe, A.,
Shaffer, T. and Hood, L.
Direct Submission
Submitted (25-APR-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
3 (bases 1 to 190565)
Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,
Pate, D. and Hood, L.
Direct Submission
Submitted (27-MAR-2001) Multimegabase Sequencing Center, Institute
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
98105, USA
On Mar 27, 2001 this sequence version replaced gi:8247780.

Center: Multimegabase Sequencing Center
Center code: UMSC
Web site: http://chroma.mbt.washington.edu/msg_wwww
Contact: leetowen@systemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399

Location/Qualifiers
1. .190565
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q31"
/clone="RP11-325L17"
/clone_lib="RP11-325L17"
985. .1035

FEATURES
source
Location/Qualifiers
1. .190565
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q31"
/clone="RP11-325L17"
/clone_lib="RP11-325L17"
985. .1035

unsure

```
/note="low quality data"
3865. .3870
/note="low quality data"
14875. .14880
/note="low quality data"
17060. .17095
/note="low quality data"
17488. .17535
/note="low quality data"
19433
/note="low quality data"
46474. .46580
/note="low quality data"
46695. .46700
/note="low quality data"
55163. .55242
/note="low quality data"
110860. .110875
/note="low quality data"
113613. .113617
/note="low quality data"
116230. .116782
/note="Single clone coverage. PCR and transposons tried."
117677. .117680
/note="low quality data"
117730. .117748
/note="low quality data"
118505. .118736
/note="low quality data"
144847. .144850
/note="low quality data"
152067. .152071
/note="low quality data"
158953. .158990
/note="low quality data"
177622. .177705
/note="low quality data"

ORIGIN
Query Match      86.5%; Score 549; DB 9; Length 190565;
Best Local Similarity 92.9%; Pred. No. 5.4e-162;
Matches 590; Conservative 0; Mismatches 35; Indels 10; Gaps 1;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTCTCCAGATCAAAACTGTAAGCTA 60
DB 23017 CCTTTATCTTTAACTCTTGTAAAGTTTGTCTCTCCAGATCAAAAGCTGTAAGCTA 23076

QY 61 CAAATGTTTCTTCAAAATGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 23077 CAAATAGTTCTTCAAAATGAGCACCAGATGGAGTCCATGACTAAGATCTACCATGGACCC 23136

QY 121 CTGGACCGCGCTGTAGCCATGCTCCGATGTTAATGACATTAAGGCAACCCCTCCCGAG 180
DB 23137 CTGGACCGCGCTGTAGCCATGCTCCGATGTTAATGACATTAAGGCAACCCCTCCCGAG 23196

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTTCAGCGGGAGAGTGTAGCGGT 240
DB 23197 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTTCAGCGGGAGAGTGTAGAGCAGT 23256

QY 241 CATCAGCCAACTCCCCCAACAGCAGCTTGGGTTTCTCTGTGAGAGGGGGAGCTGAGAGAC 300
DB 23257 TGTGAGCCAACTCCCCCAACAGCAGCTTGGGTTTCTCTGTGAGAGGGGGTACTGAGAGAC 23316

QY 301 AGGACTAGCTGGATTTTCTTAGGCCAACGAAAGAAATCCCTAAGCTAGCTGGGAAGTGACT 360
DB 23317 AGGACTAGCTGGATTTTCTTAGGCCAACTAAGAAATCCCTAAGCTAGCTGGGAAGTGACC 23376

QY 361 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCAGCCGACCAATCAGAGAGCTC 420
DB 23377 ACATCCACCTTTAAACACAGGGGTTGCACTTAGCTCAGCCGACCAATCAGAGAGCTC 23436

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCGCTG 480
```

```
Db 23437 ACTAAATGCTAATTAGGCAAAACACGAGGTAAGAAATAGCCAAATCATCTATCGCTG 23496
QY 481 AGACACAGCGGAGGACAGAGATCGGATATAAACCCAGGCAATTCGAGCGGCAACGG 540
DB 23497 AGACACAGCAGGAGGACAAATGATCGGATATAAACCCAGGCAATTCGAGCAGCAATGG 23556
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTTCACCTCTATTTCACTCT 600
DB 23557 CTACCTCTTTGGGTCCCTCCCTTTGTATGGGAGCTCAGT-----TTTCACTCT 23606
QY 601 ATTAAATCTTGCAACTGAAAAAAGAAAAA 635
DB 23607 ATTAAATCTTGCAACTGCAAAACAAACAAACAAA 23641

RESULT 29
CNS01DVH/c 203777 bp DNA linear PRI 04-MAY-2001
LOCUS Human chromosome 14 DNA sequence BAC C-2547L24 of library Caltech-D
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL135818
VERSION AL135818.3 GI:8217905
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203777)
AUTHORS Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., DeBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissenbach,J.
TITLE Sequencing of the human chromosome 14
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 203777)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Genoscope.

COMMENT
- Web : www.genoscope.cns.fr
- On Jun 3, 2000 this sequence version replaced gi:7406569.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Web site: http://www.genoscope.cns.fr/
Contact: Seqrefgenoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-30513
Downstream BAC (overlapping the SP6 end) : R-895M11 (AC=AL133153)
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.44x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 : 2
1 - 9 : 29
10 - 19 : 188
20 - 29 : 490
30 - 39 : 1165
40 - 49 : 5179
50 - 59 : 5946
60 - 69 : 9708
70 - 79 : 23956
80 - 89 : 61022
90 - 99 : 96092
-----
Percentage of bases with a quality value >= 40 : 99 %.
```

```
Location/Qualifiers
1..203777
/organism="Homo sapiens"

FEATURES
source
```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone_lib="C-2547L24"
/clone_lib="CalTech-D"
9045..9309
/notes="matching EMBL:R74544
Rdb:RH53560
dbSTS:ST342868
Identified using the e-PCR software (G. Schuler)"
16628..16834
/notes="matching EMBL:G20331
Rdb:RH16802
dbSTS:ST313937
Identified using the e-PCR software (G. Schuler)"
42379..42560
/notes="matching EMBL:N80119
Rdb:RH46446
dbSTS:ST339514
Identified using the e-PCR software (G. Schuler)"
145247..145418
/notes="matching EMBL:T02957
Rdb:RH53817
dbSTS:ST19329
Identified using the e-PCR software (G. Schuler)"
162745..163039
/notes="matching EMBL:R51241
Rdb:RH53819
dbSTS:ST42905
Identified using the e-PCR software (G. Schuler)"
167890..168003
/notes="matching EMBL:G38628
Rdb:RH83840
dbSTS:ST59195
Identified using the e-PCR software (G. Schuler)"

ORIGIN
Query Match      86.5%; Score 549; DB 9; Length 203777;
Best Local Similarity 92.9%; Pred. No. 5.5e-162;
Matches 590; Conservative 0; Mismatches 35; Indels 10; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTTGTGTTGTTCTCTCCAGATCAAACTGTAAACTA 60
DB 10333 CCCTTTATCTTTAACTCTTGTGTTGTTCTCTCCAGATCAAACTGTAAACTA 10274

QY 61 CAAATTTCTTTCAAATGGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 10273 CAAATAGTTCTTCAAATGGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 10214

QY 121 CTGACCGGCGTGTAGCCCATGTCGATGTTAATGACATTTGAGGACACCCCTCCGAG 180
DB 10213 CTGACCGGCGTGTAGCCCATGTCGATGTTGATGACATCGAAGGACACCCCTCCGAG 10154

QY 181 GAATCTCAACTGCACACACCTACTATGCCCAATTCAGCGGGAACGTTAGAGCGGT 240
DB 10153 GAATCTCAACTGCACACACCTACTAGCCCAATTCAGCGGGAACGTTAGAGCAGT 10094

QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCCTGTTTGGAGGGGGGAGCTGAGAGAC 300
DB 10093 TGTGAGCAACCTCCCAACAGCACTTGGGTTTTCCTGTTTGGAGGGGGGAGCTGAGAGAC 10034

QY 301 AGGACTAGTGGATTTCTAGGCAACAGAGATCCCTAGCCCTAGCTAGCTGGAGGAGTCACT 360
DB 10033 AGGACTAGTGGATTTCTAGGCAACAGAGATCCCTAGCCCTAGCTAGCTGGAGGAGTCACT 9974

QY 361 GCATCCACCTTAACTAGGCGGCTTGAACCTTACCTACACCCGACCAATCAGAGAGCTC 420
DB 9973 ACATCCACCTTAACTAGGCGGCTTGAACCTTACCTACACCCGACCAATCAGAGAGCTC 9914

QY 421 ACTAAATGCTAAATAGGCAAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCGCTG 480

```

```

DB 9913 ACTAAATGCTAAATAGGCAAAACAGGAGGTAAAGAAATAGCCAAATCATCTATCCCTG 9854
QY 481 AGAGCAGCAGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCGAGCGGCAACGG 540
DB 9853 AGAGCAGCAGCAGGAGGACATGATCGGATATAAACCCAGGCACTTCGAGCGGCAATGG 9794
QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTGTGTTTCACTCTATTCTACTCT 600
DB 9793 CTACCCCTTTGGTCCCTCCCTTTGATGGAGCTCAGT-----TTTCACTCT 9744

QY 601 ATTAATCTTGAACCTGAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 635
DB 9743 ATTAATCTTGAACCTGCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9709

RESULT 30
AC109992 112544 bp DNA linear PRI 25-OCT-2002
LOCUS Homo sapiens 3 BAC RP11-383G6 (Roswell Park Cancer Institute Human
DEFINITION BAC Library) complete sequence.
AC109992
AC109992.6 GI:21591818
HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 112544)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J.J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louleghed,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheehwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Neelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokweto,S., Ogih,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L., Quiles,M., Remy,G., Rives,M.,
Rojaas,R., Rojuboan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S.,
Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,B.,
Sonaika,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tatten,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 112544)
Worley,K.C.

```

TITLE
JOURNAL
REFERENCE
AUTHORS

Fri Feb 25 16:26:29 2005

us-09-319-156b-6.rge

TITLE
JOURNAL
Submitted (09-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (26-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (25-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
On Jun 26, 2002 this sequence version replaced gi:20976483.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and CDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

FEATURES

Source

Location/Qualifiers
1. .112544
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-383G6"
complement(46..1498)
/rpt_family="L1M4"
complement(1499..1804)
/rpt_family="AluSc"
complement(1805..2516)
/rpt_family="L1M4"

repeat_region

repeat_region

repeat_region

repeat_region
complement(2556..2621)
/rpt_family="AluJ/FLAM"
complement(2622..2786)
/rpt_family="L1M4"
complement(2792..2833)
/rpt_family="L1M4"
complement(2834..3134)
/rpt_family="AluY"
complement(3135..3392)
/rpt_family="L1M4"
5391..5415
/rpt_family="AT_rich"
complement(5509..5805)
/rpt_family="AluY"
6389..6551
/rpt_family="MIR"
6981..7109
/rpt_family="AluSg/x"
7173..7203
/rpt_family="(TTTA)n"
complement(7206..7491)
/rpt_family="AluSx"
complement(7505..7618)
/rpt_family="L1MC5"
7761..8048
/rpt_family="AluSx"
10339..10837
/rpt_family="AluSg"
11017..11061
/rpt_family="MADB1"
11176..11213
/rpt_family="AT_rich"
complement(11215..11374)
/rpt_family="U1"
complement(11387..11681)
/rpt_family="AluSx"
12173..12203
/rpt_family="AT_rich"
12313..12350
/rpt_family="A-rich"
13076..13361
/rpt_family="AluSx"
14236..14529
/rpt_family="L1M4"
complement(15115..15406)
/rpt_family="AluSx"
15454..15878
/rpt_family="L1MC5"
16207..16227
/rpt_family="AT_rich"
16946..17002
/rpt_family="AT_rich"
18059..18361
/rpt_family="AluJo"
19122..19425
/rpt_family="AluSx"
complement(19530..20155)
/rpt_family="L2"
complement(20156..20458)
/rpt_family="AluSg1"
complement(20459..20559)
/rpt_family="L2"
20687..20790
/rpt_family="MLT1J2"
21217..21344
/rpt_family="LTR33A"
complement(21391..21514)
/rpt_family="MERSA"
22051..22099
/rpt_family="L2"
22410..22545

Query Match

86.3%; Score 547.8; DB 9; Length 112544;

```
Best Local Similarity 92.4%; Pred. No. 1.2e-161;
Matches 587; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 1 CCTGTATCTTTAACTCCTCTGTTAAAGTTGTCTCTTCGAAATCAAAAAGTGTAAACTA 60
Db 34048 CCTGTATCTTTAACTCCTCTGTTAAAGTTGTCTCTTCGAAATCAAAAAGTGTAAACTA 34107

QY 61 CAAATGTTCTTCAAAATGAGACACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 34108 CAAATGTTCTTCAAAATGAGACACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 34167

QY 121 CTGACCGGCTCTAGCCCATCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 180
Db 34168 CTGACCGGCTCTAGCCCATCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 34227

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTGTAGACGGT 240
Db 34228 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTGTAGACAGT 34287

QY 241 CATCAGCCACCTCCCAACAGCAGCTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 34288 CATCAGCCACCTCCCAACAGCAGCTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 34347

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACT 360
Db 34348 AGGACTAGCTACATTTCTTAGGCCGAGTAAATCCCTAAGCCCTAGCTGGGAAGGTGACC 34407

QY 361 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCGACCAATCAGAGAGCTC 420
Db 34408 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCGCTGATCAATCAGAGAGCTC 34467

QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCCATCATCTATTGCTG 480
Db 34468 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCCATCATCTATTGCTGCTG 34526

QY 481 AGAGCAGCGGAGGAGGACGATCGGATATAAACCCAGGATTCGAGCCGCGCAACGG 540
Db 34527 ACAGCAGTGGGAGGAGGACGATCGGATGTAAACCCAGCATTCAAGCGCGCAACAG 34586

QY 541 CAACCCCTTTGGTCCCTCCCTCTGTTGATGGGCGCTCTGTTTCACTCTATTCTCT 600
Db 34587 CAACCCCTTTGGTCCCTCCCTCTGTTGATGGGAGCTCTGTTTCACTCTATTCTCTCT 34646

QY 601 ATTAATCTTGCACACTGAAAAAAGAAAAA 635
Db 34647 ATTAATCTTGCACACTGAAAAAAGAAAAA 34681

RESULT 31
AC009443
LOCUS Homo sapiens clone RP11-1L20, WORKING DRAFT SEQUENCE, 6 unordered
DEFINITION Homo sapiens clone RP11-1L20, WORKING DRAFT SEQUENCE, 6 unordered
pieces.
ACCESSION AC009443
VERSION AC009443.4 GI:11136808
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 192178)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-1L20
Unpublished
2 (bases 1 to 192178)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Headford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marris,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Testaye,S., Toriella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (22-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 10, 2000 this sequence version replaced gi:729764.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1480
Center clone name: 1.L.20
----- Summary Statistics
Sequencing vector: M13; W77815; 52% of reads
Sequencing vector: Plasmid; n/a; 48% of reads
Chemistry: Dye-primer-amerham; 2% of reads
Chemistry: Dye-terminator Big Dye; 98% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 189361 bases at least Q40
Consensus quality: 190532 bases at least Q30
Consensus quality: 191001 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 191678; sum-of-contigs
Quality coverage: 14.0 in Q20 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 7649: contig of 7649 bp in length
* 7650 7749: gap of 100 bp
* 7750 9020: contig of 1271 bp in length
* 9021 9120: gap of 100 bp
* 9121 10698: contig of 1578 bp in length
* 10699 10799: gap of 100 bp
* 10799 12208: contig of 1410 bp in length
* 12209 12309: gap of 100 bp
* 12309 46811: contig of 34503 bp in length
* 46812 46912: gap of 100 bp
* 46912 192178: contig of 145267 bp in length.
* Location/Qualifiers
* 1. 192178
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /clone="RP11-1L20"
* /clone_lib="RP11-1L Human Male BAC"
* 1. 7649
* /note="assembly_fragment"
* /note="end:SP6"
* /vector_side:left"
* 7750..9020
* /note="assembly_fragment"
* 9121..10698
* /note="assembly_fragment"
* 10799..12208
* /note="assembly_fragment"
* 12309..46811
* /note="assembly_fragment"
* 46912..192178
```

Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrum,J., Mello,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Testaye,S., Toriella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (22-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 10, 2000 this sequence version replaced gi:729764.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1480
Center clone name: 1.L.20
----- Summary Statistics
Sequencing vector: M13; W77815; 52% of reads
Sequencing vector: Plasmid; n/a; 48% of reads
Chemistry: Dye-primer-amerham; 2% of reads
Chemistry: Dye-terminator Big Dye; 98% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 189361 bases at least Q40
Consensus quality: 190532 bases at least Q30
Consensus quality: 191001 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 191678; sum-of-contigs
Quality coverage: 14.0 in Q20 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

TITLE
JOURNAL

COMMENT

FEATURES
source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

```

/Note="assembly_fragment
Clone_end:17
vector_side:right"

ORIGIN
Query Match      86.2%; Score 547.6; DB 2; Length 192178;
Best Local Similarity 93.4%; Pred. No. 1.5e-161;
Matches 594; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 1 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 179244 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 179303

QY 61 CAAATTTGTTCTTCAATGAGACACAGATGGAGTCATGAC-TAAGATCCACCGTGGACC 119
DB 179304 CAAATCGTTTCTTCAATGAGACCCAGATGCGATGCAATTAAGATCTACAGCAGACC 179363

QY 120 CTGGACCGGCTGCTAGGCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGA 179
DB 179364 CCTAGACCGGCTGCTAGGCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGA 179423

QY 180 GGAAATCTCAACTGACACACCCCTACTATGCCCAATTGAGCGGGAAGCAGTTAGAGCGG 239
DB 179424 GGAAATCTCAACGCGACACCCCTACTATGCCCAATTGAGCGGGAAGCAGTTAGAGCAG 179483

QY 240 TCATCAGCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGA 299
DB 179484 TCGTCGATCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGA 179543

QY 300 CAGGACTAGCTGATTTCTTAGGCCAAACGAGAAATCCCTTAAGCTAGCTGGGAAGGTGAC 359
DB 179544 CAGGACTAGCTGATTTCTTAGGCCAAACGAGAAATCCCTTAAGCTAGCTGGGAAGGTGAC 179603

QY 360 TGCATCCACCTCTAAACATGGGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCT 419
DB 179604 TGCATCCACCTTTAAACACGGGGCTTGCACCTAGCTACACCCGACCAATC-AGAGCT 179661

QY 420 CACTAAATGCTTAATTAGGCAAAATAGGAGCTTAAGAAATAGCCAAATCATCTATTGCT 479
DB 179662 CACTAAATGCTTAATTAGGCAAAATAGGAGCTTAAGAAATAGCCAAATCATCTATTGCT 179721

QY 480 GAGAGCACAGCGGGGACAGGATCGGATATATAACCCAGGCAATTCGAGCGGCAAG 539
DB 179722 GAGAGCACAGCAGGAGGAGGATCGGATATATAACCCAGGCAATTCGAGCGGCAAG 179781

QY 540 GCAACCCCTTTGGGTCCTCTTGTATGGGCTCTGTTTTCACCTCTATTTCACCTC 599
DB 179782 GCAACCCCTTTGGGTCCTCTTGTATGGGCTCTGTTTTCACCTCTATTTCACCTC 179841

QY 600 TATTAATCTTGAACCTGAAGAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 635
DB 179842 TATTAATCTTGAACCTGAAGAAATCAAAATCAAAATCAAAATCAAAATCA 179877

RESULT 32
CNS05TF5 192899 bp DNA linear PRI 30-APR-2001
LOCUS Human chromosome 14 DNA sequence BAC R-784H12 of library RPCI-11
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL365295
VERSION AL365295.4 GI:13928046
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 192899)
AUTHORS Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., Desbordes, V., Cruaud, C.,
Gyapay, G., Saurin, W. and Weissenbach, J.
TITLE Sequencing of the human chromosome 14
JOURNAL Unpublished

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

2 (bases 1 to 192899)
Genoscope.
Direct Submission
Submitted (30-APR-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
On May 2, 2001 this sequence version replaced gi:13449984.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 6.24x in Q20 bases; sum-of-contigs

Overall quality chart :
Range : bases

```

0 :
1 - 9 :
10 - 19 : 5
20 - 29 : 28
30 - 39 : 267
40 - 49 : 2595
50 - 59 : 4214
60 - 69 : 9526
70 - 79 : 27497
80 - 89 : 66134
90 - 99 : 82633

```

Percentage of bases with a quality value >= 40 : 99 %.

```

FEATURES
source
Location/Qualifiers
1..192899
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-784H12"
/clone_lib="RPCI-11"

```

ORIGIN

```

Query Match      86.2%; Score 547.6; DB 9; Length 192899;
Best Local Similarity 93.4%; Pred. No. 1.5e-161;
Matches 594; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 1 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 14879 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 14938

QY 61 CAAATTTGTTCTTCAATGAGACACAGATGGAGTCATGAC-TAAGATCCACCGTGGACC 119
DB 14939 CAAATCGTTTCTTCAATGAGACCCAGATGCGATGCAATTAAGATCTACAGCAGACC 14998

QY 120 CTGGACCGGCTGCTAGGCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGA 179
DB 14999 CCGTACCGGCTGCTAGGCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGA 15058

QY 180 GGAAATCTCAACTGACACACCCCTACTATGCCCAATTGAGCGGGAAGCAGTTAGAGCGG 239
DB 15059 GGAAATCTCAACGCGACACCCCTACTATGCCCAATTGAGCGGGAAGCAGTTAGAGCAG 15118

QY 240 TCATCAGCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGA 299
DB 15119 TCGTCGATCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGA 15178

QY 300 CAGGACTAGCTGATTTCTTAGGCCAAACGAGAAATCCCTTAAGCTAGCTGGGAAGGTGAC 359
DB 15179 CAGGACTAGCTGATTTCTTAGGCCAAACGAGAAATCCCTTAAGCTAGCTGGGAAGGTGAC 15238

QY 360 TGCATCCACCTCTAAACATGGGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCT 419

```

```

Db 15239 TGCATCCACCTTAAACACGGGCTTGCACTTAGCTCACCCCAACCAATC--AGAGCT 15296
QY 420 CACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCT 479
Db 15297 CACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCT 15356
QY 480 GAGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACG 539
Db 15357 GAGAGCACAGGAGGAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACG 15416
QY 540 GCAACCCCTTTGGTCCCTCCCTTTGTATGGCGCTCTGTCTTCACTTATTTCACCT 599
Db 15417 GCAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTCTTCACTTATTTCACCT 15476
QY 600 TATTAATCTTGCACTGAATAAAAAAAGAAAAA 635
Db 15477 TATTAATCTTGCACTTCAAAAAAAGAAAAA 15512

```

RESULT 33

```

AC068492 172281 bp DNA linear PRI 29-OCT-2002
LOCUS Homo sapiens BAC clone RP11-809C23 from 2, complete sequence.
DEFINITION AC068492
ACCESSION AC068492
VERSION AC068492.2 GI:14010945
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 172281)
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 172281)
Lamar, B., Abbott, A. and Ureta, M.
The sequence of Homo sapiens BAC clone RP11-809C23
Unpublished (2001)
3 (bases 1 to 172281)
Waterston, R.H.
Direct Submission
Submitted (02-MAY-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 172281)
Waterston, R.
Direct Submission
Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 172281)
Waterston, R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 172281)
Waterston, R.
Direct Submission
Submitted (29-OCT-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 9, 2001 this sequence version replaced gi:7677994.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0809C23
-----

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL

```

REFERENCE

```

AUTHORS
TITLE
JOURNAL

```

COMMENT

```

-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0809C23
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-809C23; actual end is at base position 172281 of RP11-809C23.

FEATURES	source	Location/Qualifiers
repeat_region	1..172281	/organism="Homo sapiens"
repeat_region	1161..1218	/mol_type="Genomic DNA"
repeat_region	1245..1302	/db_xref="taxon:9606"
repeat_region	1313..1407	/chromosome="2"
repeat_region	2612..2712	/map="2"
repeat_region	2868..3080	/clone="RP11-809C23"
repeat_region	3088..3175	/clone_lib="RPCI-11"
repeat_region	3872..3997	/rpt_family="MIR"
repeat_region	4265..4436	/rpt_family="MIR"
repeat_region	5913..6082	/rpt_family="MIR"
repeat_region	6124..6274	/rpt_family="MER1_type"
repeat_region	6855..7214	/rpt_family="L2"
repeat_region	7808..7956	/rpt_family="MaLR"
repeat_region	7990..8112	/rpt_family="L2"
repeat_region	8188..8540	/rpt_family="MaLR"
repeat_region	8659..8760	/rpt_family="MaLR"
repeat_region	9485..9630	/rpt_family="MER53"
repeat_region		/rpt_family="MaLR"

repeat_region 9785. .9844
/rpt_family="MaLR"
repeat_region 11223. .11294
/rpt_family="MIR"
repeat_region 12678. .12781
/rpt_family="MIR"
repeat_region 13515. .13713
/rpt_family="MIR"
repeat_region 14305. .14830
/rpt_family="L1"
repeat_region 14831. .15039
/rpt_family="MER1_type"
repeat_region 15040. .15071
/rpt_family="L1"
repeat_region 15072. .15405
/rpt_family="MaLR"
repeat_region 15421. .15498
/rpt_family="L1"
repeat_region 15499. .15814
/rpt_family="Alu"
repeat_region 15815. .16113
/rpt_family="L1"
repeat_region 16173. .16243
/rpt_family="L1"
repeat_region 16244. .16557
/rpt_family="Alu"
repeat_region 16558. .16645
/rpt_family="L1"
repeat_region 16646. .16830
/rpt_family="ERV1"
repeat_region 16817. .17794
/rpt_family="ERV1"
repeat_region 17795. .18118
/rpt_family="ERV1"
repeat_region 18119. .19275
/rpt_family="L1"
repeat_region 19258. .19362
/rpt_family="L1"
misc_feature 19363. .20255
/notes="similar to EST BE900547 (NID:g10388822)"
misc_feature 19363. .20203
/notes="similar to EST AUI32006 (NID:g10992360)"
misc_feature 19363. .20194
/notes="similar to EST BE900431 (NID:g10388587)"
misc_feature 19363. .20184
/notes="similar to EST BE733455 (NID:g10147447)"
misc_feature 19363. .20146
/notes="similar to EST AUI37312 (NID:g10997851)"
misc_feature 19363. .20106
/notes="similar to EST AUI35322 (NID:g10995861)"
misc_feature 19363. .20093
/notes="similar to EST AUI23565 (NID:g10948281)"
misc_feature 19363. .19632
/notes="similar to EST N83917 (NID:g1259542)"
misc_feature 19366. .20209
/notes="similar to EST BE300406 (NID:g9184154)"
misc_feature 19374. .20261
/notes="similar to EST BE741444 (NID:g10155436)"
misc_feature 19374. .20051
/notes="similar to EST BE275889 (NID:g9150850)"
misc_feature 19374. .20027
/notes="similar to EST BE410450 (NID:g9346900)"
misc_feature 19374. .19487
/notes="similar to EST BE736692 (NID:g10150684)"
misc_feature 19375. .19772
/notes="similar to EST BF043529 (NID:g10760584)"
misc_feature 19384. .20095
/notes="similar to EST BE903326 (NID:g10394419)"
misc_feature 19384. .20086
/notes="similar to EST BE901941 (NID:g10391598)"
misc_feature 19384. .20059
/notes="similar to EST AUI36548 (NID:g10997087)"
misc_feature 19384. .20010

misc_feature /note="similar to EST AUI36949 (NID:g10997488)"
19384. .19962
/note="similar to EST BF204034 (NID:g11097620)"
Query Match 86.2%; Score 547.4; DB 9; Length 172281;
Best Local Similarity 92.4%; Pred. No. 1.7e-161;
Matches 587; Conservative 0; Mismatches 46; Indels 2; Gaps 1;
Qy 1 CCCTGTATCTTTAACTCTTGTAGTTTGTCTCTCCAGAAATCAAACTGTAAAACTA 60
Db CCCTGTATCTTTAACTCTTGTAGTTTGTCTCTCCAGAAATCAAACTGTAAAACTA 17562
Qy 61 CAAATTCCTTCAAATGGAGACACAGATGAGTCCATGATCAAGATCCACCGTGGACCC 120
Db CAAATTCCTTCAAATGGAGACACAGATGAGTCCATGATCAAGATCCACCGTGGACCC 17622
Qy 121 CTGGACCGGCTGCTAGCCCATGCTCGGATGTTAATGACATTTGAAGCACCCTCCCGAG 180
Db CTGGACCGGCTGCTAGCCCATGCTCGGATGTTAATGACATTTGAAGCACCCTCCCGAG 17682
Qy 181 GAATCTCAACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db GAATCTCAACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 17742
Qy 17683 GAAATCTCAACTGCACAGCCCCCTACTATGCCCCAATTTAGCAGGAAGCAGTTAGAGTGT 17742
Qy 241 CATCAGCAAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
Db CATCGGCCAACCTCCCAATAGCACTTAGGTTTTCTGTTGAGAGGGGGTACTAAGAGGC 17802
Qy 301 AGGACTAGTGGATTTCTTAGGCCAACGAAGATCCCTAGCCTAGCTGGGAGAGTGACT 360
Db AGGACTAGTGGATTTCTTAGGCCAACGAAGATCCCTAGCCTAGCTGGGAGAGTGACT 17862
Qy 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGCAACCAATCAGAGAGCTC 420
Db ACATCCACCTTTAAACATGGGCTTGCAATTTAGCTCACACCCCAACCAATCAGAGAGCTC 17922
Qy 421 ACTAAATGCTAATTAGGCAAAAATAGAGGTAAGAAATAGCCAAATCATCTATGCTG 480
Db ACTAAATGCTAATTAGGCAAAAATAGAGGTAAGAAATAGCCAAATCATCTATGCTG 17982
Qy 481 AGAGCAGCGGGAGGACAGAGATCGGATATATAACCCAGGCATTCAGAGCGGCAACGG 540
Db AGAGCAGCGGGAGGAGCAATGATGGATATATAACCCAGGCATTCAGAGCGGCAATGG 18042
Qy 541 CAACCCCTTTGGGTCCTCCCTTGTATGCGGCTCTGTTTTCACCTCTATTTTCACCTCT 600
Db CTACCTCTTTGAGTCCCTCCCTTGTATGGGAG--CTGTTTTCACCTCTATTTTCACCTCT 18100
Qy 601 ATTAATCTTGAACCTGAAAAAAGAAAAA 635
Db ATTAATCTTGAACCTGCGAAAAAAGAAAAA 18135
RESULT 34
AC104163/c
LOCUS AC104163 169418 bp DNA linear PRI 31-MAY-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-66G8, complete sequence.
ACCESSION AC104163 AC073868
VERSION AC104163.2 GI:21281541
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 169418)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.,
and Haugen, E.D.
Direct Submission
JOURNAL Unpublished
TITLE Direct Submission
REFERENCE 2 (bases 1 to 169418)
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission

Submitted (05-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 169418)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Sanphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.,
and Haugen, E.D.
Direct Submission
Submitted (31-MAY-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On May 31, 2002 this sequence version replaced gi:17352427.

----- Genome Center -----
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgctgs@u.washington.edu
Drafting Center: WUGSC
----- Project Information -----
Center project name: chr-3
Center clone name: RP11-66G8 (bc0186)
----- Summary Statistics -----
Sequencing vector: unknown; 62% of reads
Sequencing vector: plasmid; L08752; 38% of reads
Chemistry: Dye-terminator ET; 66% of reads
Chemistry: Dye-terminator Big Dye; 34% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 169322 bases at least Q40
Consensus quality: 169405 bases at least Q30
Consensus quality: 169418 bases at least Q20
Insert size: 169418; sum-of-contigs
Quality coverage: 12.2x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': RP11-420K5 (UWGC:bc0651) AC104444, 57330-bp overlap
3': RP11-433D3 (UWGC:bc0429) AC104188, 73864-bp overlap

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.
ECORI BglII HindIII

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	8873	8104	8155	459	<800
6	<800	2067	2051	6382	6462

QY	481	AGAGCACAGCGGAGGACAGATCGGATATAAAACCCAGGCAATTCGAGCGGCAACGG	540
Db	160039	AGAGCACAGTGGGTGGGACAAAGATCGGATATAAAACCCAGGCAATTTGAACATGGCAATGG	159980
QY	541	CAACCCCCCTTTGGGTCCCTCCCTTTGTATGGGGCGCTCTGTTTTCACCTCTATTTTACACTCT	600
Db	159979	CACCCCTTTGGGTCTCTCCCTTTGTATGGGAGCTCTGTTTTCACCTCTATTTTACACTCT	159980
QY	601	ATTAATCTTGCACACTGAAAAAAGAAAAAAGAAAAA 634	
Db	159919	ATTAATCTTGCACACTCCAAAAAAGAAAAAAGAAAAA 159886	
RESULT 35			
AC104444		161054 bp DNA linear PRI 19-APR-2002	
LOCUS		Homo sapiens chromosome 3 clone RP11-420K5, complete sequence.	
DEFINITION		AC104444	
ACCESSION		AC104444.2	
VERSION		AC104444.2 GI:20198585	
KEYWORDS		HTG.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1 (bases 1 to 161054)	
JOURNAL		Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,	
REFERENCE		Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.	
AUTHORS		Direct Submission	
TITLE		Unpublished	
JOURNAL		2 (bases 1 to 161054)	
REFERENCE		Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.	
AUTHORS		Direct Submission	
TITLE		Submitted (11-DEC-2001) Genome Center, University of Washington,	
JOURNAL		Box 352145, Seattle, WA 98195, USA	
REFERENCE		3 (bases 1 to 161054)	
AUTHORS		Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,	
TITLE		Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.	
JOURNAL		Direct Submission	
REFERENCE		Submitted (19-APR-2002) Genome Center, University of Washington,	
AUTHORS		Box 352145, Seattle, WA 98195, USA	
TITLE		On Apr 19, 2002 this sequence version replaced gi:17488626.	
JOURNAL		----- Genome Center	
COMMENT		Center: University of Washington Genome Center	
		Center Code: UWGC	
		Web site: http://www.genome.washington.edu	
		Contact: uwgchgs@u.washington.edu	
		----- Project Information	
		Center project name: chr-3	
		Center Clone name: RP11-420K5 (bc0651)	
		----- Summary Statistics	
		Sequencing vector: plasmid; 100% of reads	
		Chemistry: Dye-terminator ET; 89% of reads	
		Chemistry: Dye-terminator Big Dye; 11% of reads	
		Assembly program: Phrap; version 0.990319	
		Consensus quality: 160810 bases at least Q40	
		Consensus quality: 161049 bases at least Q30	
		Consensus quality: 161054 bases at least Q20	
		Insert size: 161054; sum-of-contigs	
		Quality coverage: 11.0x in Q20 bases; sum-of-contigs	

		Overlapping Sequences:	
		5': RP11-241K18 (UWGC:bc0320) AC104850	
		3': Mapping in progress	

		Sequence Quality Assessment:	
		This entry has been annotated with sequence quality	
		estimates computed by the Phrap assembly program.	
		All manually edited bases have been reduced to quality zero.	
		Quality levels above 40 are expected to have less than	
		1 error in 10,000 bp.	
		Base-by-base quality values are not generally visible from the	
		GenBank flat file format but are available as part	
		of this entry's ASN.1 file	

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.

BgIII			EcoRI			HindIII		
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap
4139	4075	8696	8846	5016	4986			
2067	2050	6	<800	6382	6692			
4733	4896	4633	4582	512	<800			
8157	8220	4499	4384	449	<800			
66	<800	214	<800	8852	8947			
6300	6257	1869	1869	11649	11652			
7206	7163	66	<800	7791	7977			
10189	10145	6123	5997	2243	2336			
990	991	3677	3678	11879	11652			
2018	2050	7589	7598	11549	11652			
6744	6810	265	<800	1113	1110			
2069	2050	4453	4384	1119	1110			
784	772	1970	1952	2300	2336			
3476	3398	13599	13498	2080	2102			
1165	1150	4562	4582	6024	5884			
3662	3684	3815	3678	2278	2336			
174	<800	10	<800	1839	1862			
5486	5446	599	<800	4646	4692			
1285	1285	5326	5404	1421	1401			
4875	4896	2528	2566	3376	3400			
350	<800	10560	10460	2359	2336			
327	<800	1581	1540	1906	1862			
6737	6810	2287	2274	863	899			
173	<800	516	<800	3833	3828			

FEATURES

source
Location/Qualifiers
1. .161054
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-420K5"
/clone_lib="RPCI human BAC library 11"

Query Match 86.1%; Score 546.6; DB 9; Length 161054;
Best Local Similarity 92.4%; Pred. No. 3.1e-161;
Matches 587; Conservative 0; Mismatches 44; Indels 4; Gaps 1;

RESULT 37	CNS06C7Q/c	169234 bp	DNA	linear	PRI 28-APR-2001
LOCUS	CNS06C7Q/c	169234 bp	DNA	linear	PRI 28-APR-2001
DEFINITION	Human chromosome 14 DNA sequence BAC R-21A20 of library RPCI-11 from chromosome 14 of Homo sapiens (Human), complete sequence.				
ACCESSION	AL390799				
VERSION	AL390799.4	GI:13516219			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 169234)				
AUTHORS	Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P., Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F., Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C., Gypay, G., Saurin, W. and Weissenbach, J.				
TITLE	Sequencing of the human chromosome 14				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 169234)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (26-APR-2001) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr				
	On Apr 2, 2001 this sequence version replaced gi:13016608.				
	----- Genome Center				

Center: Genoscope / Centre National de Sequencage
 Center code: GS
 Web site: <http://www.genoscope.cns.fr/>
 Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.

----- Summary Statistics
 Assembly program: Phrap; version 2.0
 Quality coverage: 8.21x in Q20 bases; sum-of-contigs

 Overall quality chart :

Range	bases
0 - 9	5
10 - 19	24
20 - 29	84
30 - 39	1610
40 - 49	3705
50 - 59	5060
60 - 69	13612
70 - 79	43906
80 - 89	101228
90 - 99	

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES
 source

1. 169234
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosomes="14"
 /clone="R-21A20"
 /clone_lib="RPC1-11"
 136833. 137011
 /notes="matching EMBL:Z52518
 RHdb:RH31311
 dbSTS:STS28504
 Identified using the e-PCR software (G. Schuler)"
 160586. 160738
 /notes="matching EMBL:T77897
 RHdb:RH44683
 dbSTS:STS37751
 Identified using the e-PCR software (G. Schuler)"

ORIGIN

Query Match 86.0%; Score 546; DB 9; Length 169234;
 Best Local Similarity 93.2%; Pred. No. 4.8e-161;
 Matches 593; Conservative 0; Mismatches 40; Indels 3; Gaps 2;

Qy	1	CCCTGTATCTTTAACTCTTGTAACTTGTCTTCCAGAAATCAAACTGTAAACTA	60
Db	12949	CCCTGTATCTTTAACTCTTGTAACTTGTCTTCCAGAAATTTGAAGCTGTAAACTA	12890
Qy	61	CAAAATGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGAC-TAAGATCCACCGTGACC	119
Db	12889	CAAAATGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGAC-TAAGATCCACCGTGACC	12830
Qy	120	CTGGAGCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGA	179
Db	12829	CTTAGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGA	12770
Qy	180	GGAATCTCAATGTCACAAACCTTACTATGCCCCCAATTCAGCGGGAAGAGTAGAGCGG	239
Db	12769	GGAATCTCAATGTCACAAACCTTACTATGCCCCCAATTCAGCGGGAAGAGTAGAGCGG	12710
Qy	240	TCATCAGCAGACCTCCCAACAGACACTTGGTTTCTGTGTGAGAGGGGAGCTGAGAGA	299
Db	12709	TCGTGATCAACCTCCCAACAGACACTTGGTTTCTGTGTGAGAGGGGAGCTGAGAGA	12650
Qy	300	CAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCTAGCTGGGAGGTGAC	359
Db	12649	CAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCTAGCTGGGAGGTGAC	12590

Qy	360	TGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACCCCAACCAATCAGAGCT	419
Db	12589	TGCATCCACCTTTAAACACGGGCTTGCAACTTAGCTCACCCCAACCAATC-AGAGCT	12532
Qy	420	CACATAATGCTAATAGGCAAAATAGGAGTAAGAATAAGCAATCATCTATTGCTT	479
Db	12531	CACATAATGCTAATAGGCAAAATAGGAGTAAGAATAAGCAATCATCTATTGCTT	12472
Qy	480	GAGACACAGCGGGAGGAGGACAGGATCGGATATAAACCCAGGCATTCCAGCGGCAACG	539
Db	12471	GAGACACAGGAGGAGGAGGACAGGATCGGATATAAACCCAGGCATTCCAGCGGCAACG	12412
Qy	540	GCAACCCCTTTGGGTCCCTCCCTTTGTATGGGCTCTGTTTCACTTATTTCACCTC	599
Db	12411	GCAACCCCTTTGGGTCCCTCCCTTTGTATGGGCTCTGTTTCACTTATTTCACCTC	12352
Qy	600	TATTAATCTTGCACCTGAAAAAAGAAAAAAGAAAAA	635
Db	12351	TATTAATCTTGCACCTTCAAAAAAAGAAAAAAGATA	12316

RESULT 38
 BD136199
 LOCUS

DEFINITION

BD136199 1329 bp DNA linear PAT 18-SEP-2002
 Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.

ACCESSION

VERSION
 KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PD

PF

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

```
QY 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAACTA 60
Db 257 CCCTGTATCTTCAACTCTCTGTTAAAGTTTGTCTCTTCCAGAAATGGAAGCTGTAAAGCTA 316
QY 61 CAAATGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAAGATCAACCGTGGACCC 120
Db 317 CAAATAGTTCTTCAAAATGGAAACCCAGATCGAGTCCATGACTAAATCTACCGTGGACCC 376
QY 121 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 180
Db 377 CTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAG 436
QY 181 GAAATCTCAACTGCACAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 437 GAAATCTCAACTGCACAACCCCTACTACACTCCAAATTCAGTAGGAAGCAGTTAGAGCAGT 496
QY 241 CATCAGCCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGACTGAGAGAC 300
Db 497 TGTACGCCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGAC 556
QY 301 AGGACTAGCTGGATTTCCTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAAGTGTACT 360
Db 557 AGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGGGAAGTGTACC 616
QY 361 GCATCCACCTCTAAACATGGGCTTGCACACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 617 GCATCCATCTTTAAACATGGGCTTGCACACTTAGCTCAACCCGACCAATCAGAGAGCTC 676
QY 421 ACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
Db 677 ACTAAATCTTAATCAGGCAAAACAGAGGTAAAGCAATAGCCCAATCATCTATTGCGCTG 736
QY 481 AGAGCACAGCGGAGGACAAGATCGGATATAAACCAGGCAATTCAGAGCGGCAACGG 540
Db 737 AGAGCACAGCGGAGGACAAGATCGGATATAAACCAGGCAATTCAGAGCGGCAACAG 796
QY 541 CAACCCCTTTGGGTCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 797 CAACCCCTTTGGGTCCTCCCTTGTATGGAGCTCTGTTTCACTCTATTTCACCTCT 856
QY 601 ATTAATCTTGAACCTGAA 619
Db 857 ATTAATCATGCACTGCA 875
```

```
RESULT 39
AR344389
LOCUS AR344389 1329 bp mRNA linear PAT 17-AUG-2003
DEFINITION Sequence 108 from patent US 6582703.
ACCESSION AR344389
VERSION AR344389.1 GI:33740330
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
Komurian-Pradel,F., Jollivet-Reynaud,C. and Mandrand,B.
TITLE Isolated nucleotide sequences associated with multiple sclerosis or
rheumatoid arthritis and a process of detecting
JOURNAL Patent: US 6582703-A 108 24-JUN-2003;
FEATURES
source Location/Qualifiers
1..1329
/mol_type="mRNA"
/organism="unknown"
ORIGIN
Query Match 85.8%; Score 545; DB 6; Length 1329;
Best Local Similarity 92.4%; Pred. No. 4.3e-161;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAACTA 60
```

```
Db 257 CCCTGTATCTTCAACTCTCTGTTAAAGTTTGTCTCTTCCAGAAATGGAAGCTGTAAAGCTA 316
QY 61 CAAATGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAAGATCCACCGTGGACCC 120
Db 317 CAAATAGTTCTTCAAAATGGAAACCCAGATCGAGTCCATGACTAAATCTACCGTGGACCC 376
QY 121 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 180
Db 377 CTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAG 436
QY 181 GAAATCTCAACTGCACAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 437 GAAATCTCAACTGCACAACCCCTACTACACTCCAAATTCAGTAGGAAGCAGTTAGAGCAGT 496
QY 241 CATCAGCCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGACTGAGAGAC 300
Db 497 TGTACGCCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGAC 556
QY 301 AGGACTAGCTGGATTTCCTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAAGTGTACT 360
Db 557 AGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGGGAAGTGTACC 616
QY 361 GCATCCACCTCTAAACATGGGCTTGCACACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 617 GCATCCATCTTTAAACATGGGCTTGCACACTTAGCTCAACCCGACCAATCAGAGAGCTC 676
QY 421 ACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
Db 677 ACTAAATCTTAATCAGGCAAAACAGAGGTAAAGCAATAGCCCAATCATCTATTGCGCTG 736
QY 481 AGAGCACAGCGGAGGACAAGATCGGATATAAACCAGGCAATTCAGAGCGGCAACGG 540
Db 737 AGAGCACAGCGGAGGACAAGATCGGATATAAACCAGGCAATTCAGAGCGGCAACAG 796
QY 541 CAACCCCTTTGGGTCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 797 CAACCCCTTTGGGTCCTCCCTTGTATGGAGCTCTGTTTCACTCTATTTCACCTCT 856
QY 601 ATTAATCTTGAACCTGAA 619
Db 857 ATTAATCATGCACTGCA 875
```

```
RESULT 40
AX001030
LOCUS AX001030 1329 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 12 from Patent WO9902666.
ACCESSION AX001030
VERSION AX001030.1 GI:7241266
KEYWORDS
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1329)
AUTHORS Ott,C. and Bedin,F.
TITLE RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR
ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR
DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES
JOURNAL Patent: WO 9902666-A 12 21-JAN-1999;
BIO MERIEUX (FR); OTT CATHERINE (FR)
FEATURES
source Location/Qualifiers
1..1329
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 85.8%; Score 545; DB 6; Length 1329;
Best Local Similarity 92.4%; Pred. No. 4.3e-161;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAACTA 60
```

[illegible]

Search completed: February 21, 2005, 11:52:17
Job time : 2738.62 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 21, 2005, 06:25:25 ; Search time 327.546 Seconds
(without alignments)
11476.370 Million cell updates/sec

Title: US-09-319-156B-6
Perfect score: 635
Sequence: 1 cccgtatcttttaacctctt.....tgaaaaa.....635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1980s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	635	100.0	635	2 AAV43215	AAV43215 Multiple
2	635	100.0	635	2 AAX29702	Aax29702 Clone CL6
3	635	100.0	635	12 ADG14843	Adg14843 MSRV asso
4	626	98.6	635	9 ADB84397	Adb84397 MSRV-1 as
5	617.4	97.2	2030	3 AAG63826	Aag63826 Nucleotid
6	546.4	86.0	2074	6 AAD41225	Aad41225 Human EMB
7	545	85.8	1329	2 AAV43219	AAV43219 Multiple
8	545	85.8	1329	2 AAX29704	Aax29704 Clone 5M6
9	545	85.8	1329	12 ADG14849	Adg14849 MSRV asso
10	543.4	85.6	21646	11 ACN44454	Acn44454 Human gen
11	540.2	85.1	2046	8 ACC46747	Acc46747 Human dit
12	535	84.3	1329	9 ADB84403	Adb84403 MSRV-1 as
13	533.8	84.1	2946	2 AAX77526	Aax77526 Human sec
14	533.8	84.1	2946	3 AAZ59468	Aaz59468 Human sec
15	533.8	84.1	2946	10 ADC38776	Adc38776 Human cDN
16	532.8	83.9	326014	6 ABK83296	Abk83296 Human gen
17	532.8	83.9	326014	12 ADQ94981	Adq94981 Human kin
18	532.2	83.8	2781	5 AAF55630	Aaf55630 Nucleotid
19	525.8	82.8	1136	2 AAX25660	Aax25660 Human end
20	525.8	82.8	1136	3 AAA59210	Aaa59210 3' pol ge

21	522.6	82.3	2782	2 AAX25661	Aax25661 Human end
22	522.6	82.3	2782	3 AAA59211	Aaa59211 5' non co
23	522.6	82.3	2782	5 AAH20069	Aah20069 HERV-w en
24	517.4	81.5	2930	6 AAD24195	Aad24195 Human syn
25	516.2	81.3	5693	6 ABL61744	Ab161744 Colon ade
26	511.4	80.5	7582	2 AAX25665	Aax25665 Complete
27	511.4	80.5	7582	3 AAA59215	Aaa59215 Human end
28	510.6	80.4	46340	3 ABN97978	Abn97978 Human ret
29	510.6	80.4	161334	11 ACN44334	Acn44334 Human gen
30	505.8	79.7	10499	3 ABN97929	Abn97929 Human ret
31	500.8	78.9	285020	11 ACN44958	Acn44958 Human gen
32	499.6	78.7	1894	4 AAI14608	Aai14608 Probe #45
33	499.6	78.7	1894	4 ABA56337	Aba56337 Human foe
34	499.6	78.7	1894	4 AAI35980	Aai35980 Probe #46
35	499.6	78.7	1894	4 ABA45822	Aba45822 Human bre
36	499.6	78.7	1894	4 ABA25978	Aba25978 Probe #44
37	499.6	78.7	1894	4 AAK30018	Aak30018 Human bon
38	499.6	78.7	1894	4 AAK04516	Aak04516 Human bra
39	499.6	78.7	1894	4 ABS29670	Abs29670 Human liv
40	499.6	78.7	1894	5 AAI04422	Aai04422 Probe #44
41	497.2	78.3	2411	12 ABS04589	Abs04589 Human gen
42	497.2	78.3	2411	12 ADQ67414	Adq67414 Novel hum
43	495.6	78.0	3372	2 AAX25663	Aax25663 Human end
44	495.6	78.0	3372	3 AAA59213	Aaa59213 Partial p
45	493.2	77.7	2052	4 ABA08902	Aba08902 Human rec

ALIGNMENTS

RESULT 1
AAV43215
ID AAV43215 standard; cDNA; 635 BP.
XX
AC AAV43215;
XX
DT 29-DEC-1998 (first entry)
XX
DE Multiple sclerosis associated retrovirus fragment 4.
XX
KW Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW env gene; rheumatoid arthritis-associated virus; ss.
XX
OS Multiple sclerosis associated retrovirus.
XX
FH Key Location/Qualifiers
FT CDS 1..234
FT /*tag= a
FT /product= "Encodes protein AAV71067"
XX
PN WO9823755-A1.
XX
PD 04-JUN-1998.
XX
XX 26-NOV-1997; 97MO-IB001482.
XX
PR 26-NOV-1996; 96US-00756429.
XX
PA (NMR) BIO MERIEUX.
XX
PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
XX
DR WPI; 1998-322732/28.
DR P-PSDB; AAW71067.
XX
PT New nucleic acid from retroviruses - useful for diagnosis, prevention and
PT treatment of, e.g. multiple sclerosis.
XX
PS Disclosure; Page 183; 286pp; English.
XX
CC The present sequence represents a multiple sclerosis (MS) associated
CC retrovirus (MSRV) genomic fragment used in the method of the invention.

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

CC The invention provides complete or partial genomic sequences of the MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by these genes. The invention also provides antibodies raised against the polypeptides. The genomic sequences, polypeptides and antibodies are also claimed useful for diagnosing infection by MS and rheumatoid arthritis-CC associated viruses, and also for prevention and treatment of infection CC with these viruses

XX
SQ Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;
Query Match 100.0%; Score 635; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 1e-186;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
QY 61 CAAATTTGTTCTTCAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGACCC 120
DB 61 CAAATTTGTTCTTCAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGACCC 120
QY 121 CTGGACGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 121 CTGGACGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
DB 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
QY 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTTGTAGAGGGGGGACTGAGAGAC 300
DB 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTTGTAGAGGGGGGACTGAGAGAC 300
QY 301 AGACTAGCTGGATTTCTTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 301 AGACTAGCTGGATTTCTTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
QY 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTTCACCCGACCAATCAGAGAGCTC 420
DB 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTTCACACCCGACCAATCAGAGAGCTC 420
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCCTG 480
DB 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCCTG 480
QY 481 AGAGCAGCGGGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAACGG 540
DB 481 AGAGCAGCGGGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAACGG 540
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
DB 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600

RESULT 2

AA29702
ID AAX29702 standard; DNA; 635 BP.
XX
AC AAX29702;
XX
DT 17-OCT-2003 (revised)
DT 27-AUG-2003 (revised)
DT 08-JUN-1999 (first entry)
XX
DE Clone CL6-3' from MSRV-1.
XX Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW rheumatoid polyarthritis; ss.

XX Viruses.
XX OS
XX PN FR2765588-A1.
XX
XX PD 08-JAN-1999.
XX
XX PF 07-JUL-1997; 97FR-00008816.
XX
XX PR 07-JUL-1997; 97FR-00008816.
XX (INMR) BIO MERIEUX.
XX
XX WPI; 1999-098275/09.
XX P-PSDB; AAW9552.
XX
XX Nucleic acid sequences of retrovirus called MSRV-1 - associated with multiple sclerosis or rheumatoid polyarthritis.
XX
XX Claim 1; Page 36-37; 83pp; French.
XX
XX This sequence represents clone CL6-3' from a novel multiple sclerosis related virus type 1 (MSRV1). The sequence can be used in diagnostic, prophylactic or therapeutic compositions to inhibit expression of a multiple sclerosis related virus and/or virus associated with rheumatoid polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 635; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 1e-186;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
QY 61 CAAATTTGTTCTTCAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGACCC 120
DB 61 CAAATTTGTTCTTCAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGACCC 120
QY 121 CTGGACGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 121 CTGGACGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
DB 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
QY 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTTGTAGAGGGGGGACTGAGAGAC 300
DB 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTTGTAGAGGGGGGACTGAGAGAC 300
QY 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
QY 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTTCACCCGACCAATCAGAGAGCTC 420
DB 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTTCACACCCGACCAATCAGAGAGCTC 420
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCCTG 480
DB 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCCCTG 480
QY 481 AGAGCAGCGGGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAACGG 540
DB 481 AGAGCAGCGGGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAACGG 540
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
DB 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs). Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is a claimed MSRV-associated sequence whose identity cannot be
CC accurately determined. Note: The SEQ ID numbers for the sequences as
CC displayed in the main body of the patent do not match the SEQ ID numbers
CC in the sequence listing. Consequently those sequences mentioned in the
CC claims may not be the sequences the authors intended to claim.

XX Sequence 635 BP; 188 A; 167 C; 133 G; 138 T; 0 U; 9 Other;

Query Match 98.6%; Score 626; DB 9; Length 635;

Best Local Similarity 98.6%; Pred. No. 6.4e-184;

Matches 626; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCCTCTGTTAGTTTGTCTCTCCAGAACTCAAAACTGTAAACTA 60
DB 1 CCTGTATCTTTAACTCCTCTGTTAGTTTGTCTCTCCAGAACTCAAAACTGTAAACTA 60
QY 61 CAAATTTGTTCTTCAAAATGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB 61 CAAATTTGTTCTTCAAAATGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCN 120
QY 121 CTGGACGGGCTGTAGCCATGCTCCGATGTTATGCAATTGAAGGACCCCTCCCGAG 180
DB 121 CTGGACGGGCTGTAGCCATGCTCCGATGTTATGCAATTGAAGGACCCCTCCCGAN 180
QY 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGACGGT 240
DB 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGACGGN 240
QY 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTGTGAGAGGGGGAGCTGAGAGAC 300
DB 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTGTGAGAGGGGGAGCTGAGAGAN 300
QY 301 AGGACTAGCTGGATTCTTAGGCCAACGAAGATCCCTAAGCTAGCTGGGAGGTGACT 360
DB 301 AGGACTAGCTGGATTCTTAGGCCAACGAAGATCCCTAAGCTAGCTGGGAGGTGACN 360
QY 361 GCATCCACCTCTAAACATGGGGTTCGCACTTAGCTCACACCGACCAATCAGAGAGCTC 420
DB 361 GCATCCACCTCTAAACATGGGGTTCGCACTTAGCTCACACCGACCAATCAGAGAGCTN 420
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTTG 480
DB 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTTN 480
QY 481 AGAGCAGCGGAGGAGCAGGATCGGATATAAACCCAGGATTCGAGCGGCAACGG 540
DB 481 AGAGCAGCGGAGGAGCAGGATCGGATATAAACCCAGGATTCGAGCGGCAACGN 540
QY 541 CAACCCCTTTGGTCCCTCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 541 CAACCCCTTTGGTCCCTCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCN 600
QY 601 ATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 635
DB 601 ATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 635

RESULT 5

AAA63826

ID AAA63826 standard; DNA; 2030 BP.

XX AAA63826;

AC AAA63826;

XX 06-AUG-2003 (revised)

DT 04-DEC-2000 (first entry)

XX Nucleotide sequence of the MSRV-1 3' env and LTR regions.

DB MSRV-1; pol region; long terminal repeat; LTR; RUS region; retrovirus;

XX ss.

XX Multiple sclerosis associated retrovirus.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

1396 CCCTGTATCTTTAACTCTCTTTAAAGTTGTCTCTCAGAAATTAAGCTGTAAAGCTTA 1455
61 CAAATTTGTTCTTCAATAGGACACGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
1456 CAGATGGTCTTCAAAATGGAACCCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 1515
121 CTGGACCGGCTCTAGCCCATCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG 180
1516 CTGGACCGGCTCTAGCCCATCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG 1575
181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTAAGAGCGGT 240
1576 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTAAGAGCGGT 1635
241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 300
1636 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 1695
301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
1696 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 1755
361 GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
1756 GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 1815
421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 480
1816 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 1875
481 AGAGCAGCGGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCCCGGCAACGG 540
1876 AGAGCAGCGGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCCCGGCAACGG 1935
541 CAACCCCTCTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
1936 CAACCCCTCTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 1995
601 ATTAATCTTGCACCTGAAAAAATAAAAAA 635
1996 ATTAATCTTGCACCTGAAAAAATAAAAAA 2030

```

RESULT 6

AAD41225
ID AAD41225 standard; cDNA; 2074 BP.

AC AAD41225;

XX 30-OCT-2002 (first entry)

XX Human EMBRY-2 cDNA.

Human; embryogenesis associated protein; AIDS; reproductive disorder; infertility; endometriosis; endometrial tumour; inflammatory disorder; autoimmune disorder; acquired immune deficiency syndrome; transgenic; ovarian tumour; contact dermatitis; placenta disorder; preclampsia; EMBRY-2; allergy; gene therapy; Gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 64..1692

FT /*tag= a

FT /product= "EMBRY-2 protein"

FT sig_peptide 64..123

FT /*tag= b

FT mat_peptide 124..1689

FT /*tag= c

FT /product= "Mature EMBRY-2 protein"

XX W0200248362-A2.

XX 20-JUN-2002.
XX 14-NOV-2001; 2001WO-US043956.
XX 15-NOV-2000; 2000US-0249407P.
XX (INCY-) INCYTE GENOMICS INC.
XX Ramkumar J, Arvizu C;
XX WPI; 2002-537629/57.
XX P-PSDB; AAE25054.
XX New polypeptides of human embryogenesis associated proteins for screening
XX modulators useful for treating or preventing disorders e.g.
XX endometriosis, infertility, allergy, preclampsia.
XX Claim 59; Page 96-97; 97pp; English.
XX The invention relates to human embryogenesis associated proteins (EMBRY)
XX and nucleic acid molecules encoding such proteins. EMBRY sequences are
XX useful for screening modulators useful for treating or preventing
XX disorders associated with abnormal expression of EMBRY. The disorders
XX treated include reproductive disorders such as infertility.
XX endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
XX disorder such as acquired immune deficiency syndrome (AIDS), allergies,
XX contact dermatitis; disorders of the placenta such as preclampsia,
XX abruptio placentae etc. Sequences of the invention are also useful for
XX analysing a proteome of a tissue or a cell type. EMBRY proteins are
XX useful as immunogens for preparing antibodies. Polynucleotides of the
XX invention are useful for creating knockin humanised animals or transgenic
XX animals to model human diseases. They are also used in gene therapy. The
XX present sequence is human EMBRY-2 cDNA
XX Sequence 2074 BP; 583 A; 567 C; 392 G; 531 T; 0 U; 1 Other;

Query Match 86.0%; Score 546.4; DB 6; Length 2074;

Best Local Similarity 93.8%; Pred. NO. 6.7e-159;

Matches 579; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

Qy	1	CCCTGTATCTTTAACTCTCTTTAAAGTTGTCTCTCAGAAATTAAGCTGTAAAGCTTA	60
Db	1459	CCCTGTATCTTTAACTCTCTTTAAAGTTGTCTCTCAGAAATTAAGCTGTAAAGCTTA	1518
Qy	61	CAAATTTGTTCTTCAATAGGACACGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120
Db	1519	CAAATTTGTTCTTCAATAGGACACGATGGAGTCCATGACTAAGATCCACCGTGGACCC	1578
Qy	121	CTGGACCGGCTCTAGCCCATCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG	180
Db	1579	CTGGACCGGCTCTAGCCCATCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG	1638
Qy	181	GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTAAGAGCGGT	240
Db	1639	GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTAAGAGCGGT	1698
Qy	241	CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC	300
Db	1699	CGTCAGTCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC	1758
Qy	301	AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT	360
Db	1759	AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT	1818
Qy	361	GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC	420
Db	1819	GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC	1878
Qy	421	ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG	480
Db	1879	ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG	1938

	Best Local Similarity Matches	92.4% Conservative	Pred. NO. 1.5e-158; 0; Mismatches	Indels	Gaps
2y	1	CCCTGTAATTTTAAACCTCTCTGTTAAGTTTGTCTCTTCCAGAACTCAAACCTGTTAAACTTA	60		
Db	257	CCCTGTAATTTTAAACCTCTCTGTTAAGTTTGTCTCTTCCAGAACTGAAAGCTGTTAAAGCTA	316		
2y	61	CAAAATGTTCTTTCAAAATGAGACACAGATGGAGTCCATGACTAGATCCACCGTGGACCC	120		
Db	317	CAAAATAGTTCCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAGATCCACCGTGGACCC	376		
2y	121	ATCGACATCGCTCTGATAGCCATGCTCCGATGTTAATGACTGATTGAAGGACACCCCTCCCGAG	180		

DB		377	CTGACGGCGCCCTCCTAGACATCAAGTTCCTGTTGAGAGGGGTGGACTGAGCGGT	240
QY		181	GAAATCTCAA CTGCA CAA GCCCTACTATGCCCCAATT CAGCGGGAAGCAGTTAGAGCGGT	240
Db		437	GAAATCTCAA CTGCA CAACCCCTACTACACT CCAATTCAGTAGGHAAGCAGTTAGGNCAGT	496
QY		241	CATCAGGCAACCTCCCCAACAGCACCTTGCGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC	300
Db		497	TGTCAGGCAACCTCCCCNACAGTACTTGCGGTTTTCTGTTGAGAGGGTGGACTGAGAGAC	556
QY		301	AGGACTAGCTGGATTTCTTAGGCCAAGAGAATCCCTTAAGCTAGCTGGGAAGTGACT	360
Db		557	AGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCNNAAGCTTANTCTGGGNAGTGACC	616
QY		361	GCATCCACCTCTAAA CATGGGGGCTTGCAACTTAGCTCACACCGACCAATCAGAGAGCTC	420
Db		617	GCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCGACCAATCAGAGAGCTC	676
QY		421	ACTAAATGCTTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGTCCTG	480
Db		677	ACTAAATGCTTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCCAATCATCTATTGTCCTG	736
QY		481	AGAGCACAGCGGGAGGACAAGGATCGGGATATAAA CCCAGGATTTGAGCGCGGCAACGG	540
Db		737	AGAGCACAGCGGGAAGACAAGGATTTGGGATATAAA ACTCAGGCGATTTCAAGCCAGCAACAG	796

[illegible]

PA (INMR) BIO MERIEUX.
 XX
 DR WPI; 1999-098275/09.
 DR P-PSDB; AAW99554.
 XX
 PT Nucleic acid sequences of retrovirus called MSRV-1 - associated with
 PT multiple sclerosis or rheumatoid polyarthritis.
 XX
 PS Claim 1; Page 39-40; 83pp; French.
 XX
 CC This sequence represents clone 5M6 from a novel multiple sclerosis
 CC related virus type 1 (MSRV1). The sequence can be used in diagnostic,
 CC prophylactic or therapeutic compositions to inhibit expression of a
 CC multiple sclerosis related virus and/or virus associated with rheumatoid
 CC polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
 CC 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;
 Query Match 85.8%; Score 545; DB 2; Length 1329;
 Best Local Similarity 92.4%; Pred. No. 1.5e-158;
 Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 1 CCCTGTATCTTTAACTCCTCTGTAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
 DB 257 CCCTGTATCTTCACTCTCTGTAGTTTGTCTTCCAGAAATGAAGCTGTAAAGCTA 316
 QY 61 CAAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
 DB 317 CAAATAGTTCTTCAATGGAAACCCAGATGCAGTCCATGACTAAATCTACCGTGACCC 376
 QY 121 CTGACCGGCTGCTAGCCATGCTCGATGTTTAATGACATTAAGACATCAAGCCCTCCCGAG 180
 DB 377 CTGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTAAGACATCAAGCCCTCCCGAG 436
 QY 181 GAAATCTCAACTGACAAACCCCTACTATGCCCAATTCAGGGGAAACAGTGTAGAGCGT 240
 DB 437 GAAATCTCAACTGACAAACCCCTACTATGCCCAATTCAGTAGGAACAGTGTAGAGCGT 496
 QY 241 CATCAGCCAACTCCCAACAGCACTTTGCTTCCAGTCCATGACTAAGATCCACCGTGACCC 300
 DB 497 TGTGACCGGCTGCTAGCCATGCTCGATGTTTAATGACATTAAGACATCAAGCCCTCCCGAG 556
 QY 301 AGGACTAGCTGGATTTCTTAGGGCAACGAAGAATCCCTAAGCTTAGCTGGAGAGTCACT 360
 DB 557 AGGACTAGCTGGATTTCTTAGGGTGAATAAGAATCCCNAAAGCCTTANCTGGGAAGGTGACC 616
 QY 361 GCATCCACCTCTTAAACATGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
 DB 617 GCATCCACCTCTTAAACATGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 676
 QY 421 ACTAAATGCTTAATAGGCAAAATAGAGGTAAAGAAATAGCAATCACTATTGCTGT 480
 DB 677 ACTAAATGCTTAATAGGCAAAATAGAGGTAAAGCAATAGCAATCACTATTGCTGT 736
 QY 481 AGACACAGCGGGAGGACAGAGATCGGATATTAACCCAGGCAATTCGAGCCGCAACGG 540
 DB 737 AGACACAGCGGGAGGACAGAGATCGGATATTAACCCAGGCAATTCGAGCCGCAACAG 796
 QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTTATGGGGCTCTGTTTCACTTATTTCACTCT 600
 DB 797 CAACCCCTTTGGGTCCCTCCCTTTGTTATGGGGAGCTCTGTTTCACTTATTTCACTCT 856
 QY 601 ATTAATCTTGCACACTGAA 619
 DB 857 ATTAATCATGCAACTGCA 875
 RESULT 9
 ADG14849
 ID ADG14849 standard; cDNA; 1329 BP.
 XX
 AC ADG14849;

XX
 DT 26-FEB-2004 (first entry)
 XX
 DE MSRV associated cDNA #11.
 XX
 KW es; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
 XX
 OS Multiple sclerosis associated retrovirus.
 XX
 PN US2003198647-A1.
 XX
 PD 23-OCT-2003.
 XX
 PF 03-APR-2002; 2002US-00114104.
 XX
 PR 26-NOV-1996; 96US-00756429.
 PR 26-NOV-1997; 97US-00979847.
 XX
 PA (INMR) BIO MERIEUX.
 XX
 PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
 PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
 XX
 WPI; 2004-032461/03.
 XX
 CC New isolated nucleic acid and their fragments having the pol gene of a
 CC retrovirus, useful for diagnosing, preventing and/or treating multiple
 CC sclerosis and/or rheumatoid arthritis.
 XX
 PS Disclosure; SEQ ID NO 108; 193pp; English.
 XX
 CC The invention relates to an isolated nucleic acid which comprises the pol
 CC gene of a retrovirus associated with multiple sclerosis or rheumatoid
 CC arthritis. The methods and compositions of the present invention are
 CC useful for diagnosing, preventing and/or treating multiple sclerosis
 CC and/or rheumatoid arthritis. The present sequence is used in the
 CC exemplification of the invention.
 XX
 SQ Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;
 Query Match 85.8%; Score 545; DB 12; Length 1329;
 Best Local Similarity 92.4%; Pred. No. 1.5e-158;
 Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 1 CCCTGTATCTTTAACTCCTCTGTAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
 DB 257 CCCTGTATCTTCACTCTCTGTAGTTTGTCTTCCAGAAATGAAGCTGTAAAGCTA 316
 QY 61 CAAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
 DB 317 CAAATAGTTCTTCAATGGAAACCCAGATGCAGTCCATGACTAAATCTACCGTGACCC 376
 QY 121 CTGACCGGCTGCTAGCCATGCTCGATGTTTAATGACATTAAGACATCAAGCCCTCCCGAG 180
 DB 377 CTGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTAAGACATCAAGCCCTCCCGAG 436
 QY 181 GAAATCTCAACTGACAAACCCCTACTATGCCCAATTCAGGGGAAACAGTGTAGAGCGT 240
 DB 437 GAAATCTCAACTGACAAACCCCTACTATGCCCAATTCAGTAGGAACAGTGTAGAGCGT 496
 QY 241 CATCAGCCAACTCCCAACAGCACTTTGCTTCCAGTCCATGACTAAGATCCACCGTGACCC 300
 DB 497 TGTGACCGGCTGCTAGCCATGCTCGATGTTTAATGACATTAAGACATCAAGCCCTCCCGAG 556
 QY 301 AGGACTAGCTGGATTTCTTAGGGCAACGAAGAATCCCTAAGCTTAGCTGGAGAGTCACT 360
 DB 557 AGGACTAGCTGGATTTCTTAGGGTGAATAAGAATCCCNAAAGCCTTANCTGGGAAGGTGACC 616
 QY 361 GCATCCACCTCTTAAACATGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
 DB 617 GCATCCACCTCTTAAACATGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 676
 QY 421 ACTAAATGCTTAATAGGCAAAATAGAGGTAAAGAAATAGCAATCACTATTGCTGT 480
 DB 677 ACTAAATGCTTAATAGGCAAAATAGAGGTAAAGCAATAGCAATCACTATTGCTGT 736
 QY 481 AGACACAGCGGGAGGACAGAGATCGGATATTAACCCAGGCAATTCGAGCCGCAACGG 540
 DB 737 AGACACAGCGGGAGGACAGAGATCGGATATTAACCCAGGCAATTCGAGCCGCAACAG 796
 QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTTATGGGGCTCTGTTTCACTTATTTCACTCT 600
 DB 797 CAACCCCTTTGGGTCCCTCCCTTTGTTATGGGGAGCTCTGTTTCACTTATTTCACTCT 856
 QY 601 ATTAATCTTGCACACTGAA 619
 DB 857 ATTAATCATGCAACTGCA 875
 RESULT 9
 ADG14849
 ID ADG14849 standard; cDNA; 1329 BP.
 XX
 AC ADG14849;

Db 677 ACTAAATCTTAATCAGGCAAAAACAGGAGTAAAGCAATATACCAATCACTATTGGCTG 736
 QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCGAGCCGCAACGG 540
 Db 737 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCGAGCCGCAACAG 796
 QY 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGCTCTGTTTCACTCTATTCACTCT 600
 Db 797 CAACCCCTTTGGGTCCCTCCCTTTGATGGGCTCTGTTTCACTCTATTCACTCT 856
 QY 601 ATTAATCTTGAACATGAA 619
 Db 857 ATTAATCTTGAACATGCA 875

RESULT 10
 ACN44454
 ID ACN44454 standard; DNA; 21646 BP.

AC ACN44454;
 XX
 XX 18-NOV-2004 (first entry)
 XX
 XX Human genomic sequence hCG1748307.

XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.
 XX WO2003073826-A2.

XX 12-SEP-2003.
 XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 910; Opp; English.

CC The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published

XX Sequence 21646 BP; 5145 A; 5617 C; 5218 G; 5646 T; 0 U; 20 Other;

Query Match 85.6%; Score 543.4; DB 11; Length 21646;
 Best Local Similarity 94.8%; Pred. No. 1.8e-157;
 Matches 562; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTGTTAAAGTTGCTCTCCAGATCAAACTGTAAACTA 60
 Db 19571 CCCTGTATCTTTAACTCTCTGTTAAAGTTGCTCTCCAGATCAAACTGTAAACTA 19630

QY 61 CAAATTTCTTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 19631 CAAATGGTTCTTCAAAATGGAGCCCCAGATGCAGTCCATGACTAAGATCCACCAACAGACCC 19690
 QY 121 CTGACCGCGCTGCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCGAG 180
 Db 19691 CTGACCGAGCTGCTAGCCCATGCTCCGATGTTAATGACATCGAAGGACACCCCCCGAG 19750
 QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 240
 Db 19751 GAAATCTCAACTGTGCAACCCCTACTATGCCCAATTCAGCGGAGAGCAGTTACAGCGGT 19810
 QY 241 CATGAGCCAACTCCCAACAGCACTTGGGTTTTCTCTGAGAGGGGGGAGTACAGAGAC 300
 Db 19811 CATGAGCCAACTCCCAACAGCACTTGGGTTTTCTCTGAGAGGGGGGAGTACAGAGAC 19870
 QY 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 360
 Db 19871 AAGACTAGCTGGATTTCTTAGGCCGACTAAGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 19930
 QY 361 GCATCCACTCTAAACATGGGCTTGCACCTTAGCTCACACCCGACCAATCAGAGAGCTC 420
 Db 19931 GCATCCACTCTTAAACATGGGCTTGCACCTTAGCTCACACCCGACCAATCAGAGAGCTC 19990
 QY 421 ACTAAATGCTAAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCTG 480
 Db 19991 ACTAAATGCTAAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCTG 20050
 QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCAGAGCCGCAACGG 540
 Db 20051 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCAGAGCCGCAATGG 20110
 QY 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGCTCTGTTTCACTCTATT 593
 Db 20111 CAACCCCTTTGGGTCCCTCCCTTTGATGGGCTCTGTTTCACTCTATT 20163

RESULT 11
 ACC46747

ID ACC46747 standard; cDNA; 2046 BP.

XX AC ACC46747;

XX 02-JUN-2003 (first entry)

XX Human dithp growth/development-associated protein-encoding cDNA.

XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging; growth;
 KW development; gene; ss.

XX Homo sapiens.

XX WO200297031-A2.

PD 05-DEC-2002.

PF 27-MAR-2002; 2002WO-US010056.

XX 28-MAR-2001; 2001US-0279619P.

PR 29-MAR-2001; 2001US-0280067P.

PR 29-MAR-2001; 2001US-0280069P.

PR 16-MAY-2001; 2001US-0291280P.

PR 17-MAY-2001; 2001US-0291829P.

PR 17-MAY-2001; 2001US-0291849P.

PR 19-JUN-2001; 2001US-0299428P.

PR 20-JUN-2001; 2001US-0299776P.

PR 20-JUN-2001; 2001US-0300001P.
XX (INCY-) INCYTE GENOMICS INC.
PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
DR P-PSDB; ABR41810.
XX Novel human diagnostic and therapeutic polypeptide useful for identifying
XX test compound which specifically binds to a polypeptide encoded by human
XX diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX Claim 2; SEQ ID NO 668; 591pp; English.
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a dithp cDNA encoding a DITHP protein
CC which is associated with growth and development. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2046 BP; 587 A; 494 C; 391 G; 566 T; 0 U; 8 Other;
Query Match 85.1%; Score 540.2; DB 8; Length 2046;
Best Local Similarity 94.4%; Pred. No. 5.6e-157;
Matches 560; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 1 CCTGTATCTTTAACTCTTGTAGTTTGTCTCTTCCAGATCAAACTGTAACCTA 60
DB 571 CTCTGTATCTTTAACTCTTGTAGTTTGTCTCTTCCAGATCAAACTGTAACCTA 630
QY 61 CAATATCTTCTTCAATGGAGCCAGATGAGTCCATGATCAATGATCCACCGTGACCC 120
DB 631 CAATATCTTCTTCAATGGAGCCAGATGAGTCCATGATCAATGATCCACCGTGACCC 690
QY 121 CTGACCGGCGCTGTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCCACCCCTCCGAG 180
DB 691 CTGACCGGCGCTGTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCCACCCCTCCGAG 750
QY 181 GAAATCTCACTGACCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 751 GAAATCTCACTGACCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 810
QY 241 CATCAGCCACCTCCCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGGGACTGAGAGAC 300

DB 811 CATCAGCCACCTCCCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGCACTGAGAGAC 870
QY 301 AGGACTAGCTGGATTTCTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 871 AAGACTAGCTGGATTTCTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 930
QY 361 GCATCCACCTCTAAACATATGGGGCTTGGCAACTTTAGCTTCACACCGGACCAATCAGAGAGCTC 420
DB 931 GCATCCACCTCTTAAACATATGGGGCTTGGCAACTTTAGCTTCACACCGGACCAATCAGAGAGCTC 990
QY 421 ACTAAATATGCTAATTTAGGCAAAAATAGGAGGTAAGAAATAGCCATCATCTATTGCGCTG 480
DB 991 ACTAAATATGCTAATTTAGGCAAAAATAGGAGGTAAGAAATAGCCATCATCTTTGCGCTG 1050
QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCCAGGCACTTCGAGCCGCGCAACGG 540
DB 1051 AGAGCACAGCAGAGGAGGACNAGATCGGGATATAAACCCGCGGCACTTCAGCCGCGCAATGG 1110
QY 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGCGCTCTGTTTCACTCTATT 593
DB 1111 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGAGCTCTGTTTCACTCTATT 1163
RESULT 12
ADB84403
ID ADB84403 standard; DNA; 1329 BP.
XX ADB84403;
AC ADB84403;
DT 04-DEC-2003 (first entry)
XX MSRV-1 associated DNA sequence #21.
XX ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
XX reverse transcriptase; ribonuclease H.
XX Unidentified;
XX US2003039664-A1.
XX 27-FEB-2003.
XX 26-NOV-1997; 97US-00979847.
XX 26-NOV-1996; 96US-00756429.
XX (PERR/) PERRON H.
XX (BESE/) BESEME F.
XX (BEDI/) BEDIN F.
XX (PARA/) PARANHOS-BACCALA G.
XX (KOMU/) KOMURIAN-PRADEL F.
XX (JOLI/) JOLIVET-REYNAUD C.
XX (MAND/) MANDRAND B.
XX (GARS/) GARSON J A.
XX (TUKU/) TUKE P W.
XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
PI WPI; 2003-512253/48.
DR New isolated or purified nucleic acid associated with multiple sclerosis
XX and/or rheumatoid arthritis, useful for detecting a virus associated with
XX multiple sclerosis or rheumatoid arthritis in a biological sample.
XX Claim 31; Page 80; 193pp; English.
XX The invention relates to an isolated or purified nucleic acid from a
XX virus associated with multiple sclerosis and/or rheumatoid arthritis,
XX multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
XX pol, gag or reverse transcriptase genes (or their fragments) encoding the
XX proteins or defined peptides (including immunodominant peptides,

CC antigenic peptides or conserved motifs). Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is a claimed MSRV-associated sequence whose identity cannot be
CC accurately determined. Note: The SEQ ID numbers for the sequences as
CC displayed in the main body of the patent do not match the SEQ ID numbers
CC in the sequence listing. Consequently those sequences mentioned in the
CC claims may not be the sequences the authors intended to claim.
XX
SQ Sequence 1329 BP; 339 A; 346 C; 282 G; 332 T; 0 U; 30 Other;

Query Match 84.3%; Score 535; DB 9; Length 1329;
Best Local Similarity 90.8%; Pred. No. 1.9e-155;
Matches 562; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60
DB |||
QY 61 CAAATTGTTCTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB |||
QY 317 CAAATAGTTCTCAATGGAAACCCAGATGCAGTCCATGACTANAACTACCGTGACCC 376
DB |||
QY 121 CTGACCGGCTCTAGCCATGCTCCGATGTTAATGACATTAAGAGCAACCCCTCCCGAG 180
DB |||
QY 377 CTGGACGGGCTCTAGACTACTCTGTATGTTAATGACATTAAGAGCAACCCCTCCCGAG 436
DB |||
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGAGTGTAGCGGT 240
DB |||
QY 437 GAAATCTCAACTGCACAAACCCCTACTACATCTCCAAATTCAGTAGNAGCAGTTAGACGT 496
DB |||
QY 241 CATCAGCAACCTCCCAACAGACACTTGGTTTCTGTGTGAGAGGGGGGACTGAGAGAC 300
DB |||
QY 497 TGTGAGCAACCTCCCAACAGTACTTGGGTTTCTGTGTGAGNGGGTGGACTGAGAGAC 556
DB |||
QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCTTAAGCCTAGCTGGGAAGGTGACT 360
DB |||
QY 557 AGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCNAAAGCCNANCTGGGAAGGTGACC 616
DB |||
QY 361 GCATCCACCTCTAAACATGGGGTTGCAACTTAGCTCACACCCGACCAATCAGAGAGTTC 420
DB |||
QY 617 GCATCCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCCNAGCAATCAGAGAGTTC 676
DB |||
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTTG 480
DB |||
QY 677 ACTAAATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCNAATCATCTATTGCTTG 736
DB |||
QY 481 AGAGCAGCGGGAGGAGCAAGGATCGGATATAAAACCCAGGCAATTCGAGCCGGCAACGG 540
DB |||
QY 737 AGAGCAGCGGGAGGAGCAAGGATCGGATATAAACTAGGATATAAACTCAGGCTTCAAGCCAGCAACAG 796
DB |||
QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGGGCTCTGTTTCACTCTATTTCACCTCT 600
DB |||
QY 797 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 856
DB |||
QY 601 ATTAATCTTGCACCTGAA 619
DB |||
QY 857 ATTAATCATGCAACTGCA 875
DB |||

AAAX77526

ID AAX77526 standard; cDNA; 2946 BP.

XX AC

XX AAX77526;

XX AC

XX 10-AUG-1999 (first entry)

XX DT

XX Human secreted protein AJ172_2 cDNA.

XX DE

XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW cell proliferation; cell differentiation; suppressor; tumour inhibition;
KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
KW cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.

XX OS

XX Homo sapiens.

XX XX

XX WO9926972-A1.

XX PN

XX 03-JUN-1999.

XX PD

XX 17-NOV-1998; 98WO-US024614.

XX PF

XX 21-NOV-1997; 97US-00976110.

XX PR

XX 18-MAY-1998; 98US-00080478.

XX PR

XX 20-OCT-1998; 98US-00175928.

XX XX

XX (GEMY) GENETICS INST INC.

XX XX

XX Jacobs K, McCoey JM, Lavallie ER, Collins-Racie LA, Evans C;

XX PI

XX Merberg D, Treacy M;

XX XX

XX WPI; 1999-357813/30.

XX DR

XX P-PSDB; AAY08622.

XX DR

XX New polynucleotides encoding secreted proteins.

XX PT

XX XX

XX Claim 13a; Page 100-101; 142pp; English.

XX PS

XX This invention describes novel human secreted proteins encoded by
CC polynucleotides isolated from human adult testes, adult brain, adult
CC blood or adult placenta, or murine adult bone marrow or thymus CDNA
CC libraries. The products of the invention are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful for
CC gene therapy

XX SQ

Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

Query Match 84.1%; Score 533.8; DB 2; Length 2946;

Best Local Similarity 91.5%; Pred. No. 6.6e-155;

Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60

DB |||

DB 2323 CCTGTATCTTTAACTCTTGTAACTTTGTCTTCCAGAACTCGAAGCTGTAAACTA 2382

QY 61 CAAATGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120

DB |||

DB 2383 -----CAAAATGGAGCCCAAGATGCGAGTCCAAAGACTAAGATCTTACCGGAGACCC 2430

QY 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 180

DB |||

DB 2431 CTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATCAAAAGGACCCCTCTCTGAG 2490

181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 2491 GAAATCTCAAGTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 2550
 241 CATCAGCAACCTCCCAACAGCAGCACTTGGGTTTCTCTTTGAGAGGGGAGCTGAGAGAC 300
 2551 CGTCGGCAACCTCCCAACAGCAGCACTTGGGTTTCTCTTTGAGATGGGAGCTGAGAGAC 2610
 301 AGACTAGCTGGATTTCTTAGGCAAGCAAGCAATCCCTAAGCCTAGCTAGCTGGGAAGTGA 360
 2611 AGGACTAGCTGGATTTCTTAGGCTGACTAAGCAATCCCTAAGCCTAGCTGGGAAGTGA 2670
 361 GCATCCACCTCTAAACATGGGCTTGCACTTAGCTCAGCAGGCAATCCCAATCAGAGAGCTC 420
 2671 ACATCCACCTTTAAACAGGGGCTTGCACTTAGCTCAGCAGGCAATCCCAATCAGAGAGCTC 2730
 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
 2731 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 2790
 481 AGAGCAGCGGAGGAGCAAGGATCGGGATATAACCCAGGCAATTCGAGCCGGCAACGG 540
 2791 AGAGCAGCAGGAGGAGCAATGATCGGGATATAACCCAGGCAATTCGAGCCGGCAACGG 2850
 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGCTCTCTTTTCACTCTATTTCACCTCT 600
 2851 CAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 2910
 601 ATTAATCTTTGCAACTGCAAAAAAAGGAGTAAAGAAATAGCAATCATCTATTGCTG 635
 2911 ATTAATCTTTGCAACTGCAAAAAAAGGAGTAAAGAAATAGCAATCATCTATTGCTG 2945

RESULT 14
 AAZ59468
 ID AAZ59468 standard; cDNA; 2946 BP.
 XX AAZ59468;
 XX
 11-APR-2000 (first entry)
 DE Human secreted protein AJ172_2 polynucleotide sequence.
 XX
 KW Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
 KW placental pathology; metastasis inhibition; nutritional activity;
 KW immune stimulator; haematopoiesis regulator; tissue growth;
 KW tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
 KW gene therapy; ss.
 XX Homo sapiens.
 OS
 XX WO9960020-A1.
 PN
 XX 25-NOV-1999.
 PD
 XX
 PF 17-MAY-1999; 99WO-US010915.
 XX
 PR 18-MAY-1998; 98US-00080478.
 PR 20-OCT-1998; 98US-00175928.
 XX
 PA (GEM) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Mi S, Treacy M;
 XX
 DR WPI; 2000-116311/10.
 DR P-PSDB; AAY67313.
 XX
 PT New polynucleotides encoding secreted cDNA libraries, used to develop
 PT products for the diagnosis and treatment of neoplastic disease.
 XX
 PS Claim 14; Page 107-108; 149pp; English.

XX This is the human secreted protein AJ172_2 nucleotide sequence, obtained
 CC from a human adult testes cDNA library. The invention relates to secreted
 CC human and murine proteins. The polynucleotides and proteins are predicted
 CC to have biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals. Detection of the levels of the proteins can be used for the
 CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
 CC which modulate the expression or function of the proteins may be used for
 CC treating a neoplastic disease and inhibiting metastasis. Other suggested
 CC activities include nutritional activity (e.g. in feeds), cytokine and
 CC cell proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity. The
 CC polynucleotide sequences are also stated to be useful for gene therapy
 XX

Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;
 Query Match 84.1%; Score 533.8; DB 3; Length 2946;
 Best Local Similarity 91.5%; Pred. No. 6.6e-155;
 Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;
 'QY 1 CCCTGTATCTTTAAACCTCCTTGTAAAGTTGTCTTCCAGAAATCAAAACTGTAACCTA 60
 DB 2323 CCCTGTATCTTTAAACCTCCTTGTAAAGTTGTCTTCCAGAAATCAAAACTGTAACCTA 2382
 QY 61 CAATTTGTTCTTCAAAATGAGACACAGATGGATCCATGATCAAGTCCACCTGGACCC 120
 DB 2383 -----CAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCCGAGACCC 2430
 QY 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
 DB 2431 CTGGACCGGCTCTAGCCCATGCTCTGATGTTTAATGACATCAAGGACACCCCTCCCGAG 2490
 QY 181 GAAATCTCAACTGCACAAACCCCTTACTATGCCCAATTCAGCGGGAAGCACTTAGAGCGGT 240
 DB 2491 GAAATCTCAAGTGCACAAACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 2550
 QY 241 CATCAGCAACCTCCCAACAGCAGCACTTGGGTTTCTCTGTTGAGGGGGAGCTGAGAGAC 300
 DB 2551 CGTCGGCAACCTCCCAACAGCAGCACTTAGGTTTCTCTGTTGAGTGGGGAGCTGAGAGAC 2610
 QY 301 AGGACTAGCTGGATTTCTTAGGCAAGCAAGCAATCCCTAAGCCTAGCTGGGAAGTGA 360
 DB 2611 AGGACTAGCTGGATTTCTTAGGCTGACTAAGCAATCCCTAAGCCTAGCTGGGAAGTGA 2670
 QY 361 GCATCCACCTCTAAACATGGGCTTGCACCTTAGCTCACAACCCGACCAATCAGAGAGCTC 420
 DB 2671 ACATCCACCTTTAAACAGCGGGCTTGCACCTTAGCTCACAACCTGACCAATCAGAGAGCTC 2730
 QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
 DB 2731 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 2790
 QY 481 AGAGCAGCGGAGGAGCAAGGATCGGGATATAACCCAGGCAATTCGAGCCGGCAACGG 540
 DB 2791 AGAGCAGCAGGAGGAGCAATGATCGGGATATAACCCAGGCTTCGAGCCGGCAACGG 2850
 QY 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 600
 DB 2851 CAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 2910
 QY 601 ATTAATCTTTGCAACTGCAAAAAAAGGAGTAAAGAAATAGCAATCATCTATTGCTG 635
 DB 2911 ATTAATCTTTGCAACTGCAAAAAAAGGAGTAAAGAAATAGCAATCATCTATTGCTG 2945
 RESULT 15
 ADC38776
 ID ADC38776 standard; cDNA; 2946 BP.

FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(56453,C)	
FT		/*tag= by	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(59300,T)	
FT		/*tag= bz	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(60998,A)	
FT		/*tag= ca	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(63013,A)	
FT		/*tag= cb	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(63293,A)	
FT		/*tag= cc	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(64181,T)	
FT		/*tag= cd	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(64191,A)	
FT		/*tag= ce	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(68992,C)	
FT		/*tag= cf	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(76246,Y)	
FT		/*tag= cg	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(77368,G)	
FT		/*tag= ch	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(78043,A)	
FT		/*tag= ci	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(79395,T)	
FT		/*tag= cj	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(79474,G)	
FT		/*tag= ck	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(79992,T)	
FT		/*tag= cl	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(81065,C)	
FT		/*tag= cm	
Query Match 83.9%; Score 532.8; DB 6; Length 326014;			
Best Local Similarity 90.5%; Pred. No. 1.3e-153;			
Matches 583; Conservative 0; Mismatches 52; Indels 9; Gaps 1;			
Qy	1	CCCTGTATCTTTAACTCCTCTGTGTAGTTGTCTCTCCAGATCAAACTGTAAACTA 60	
Db	170822	CCCTGTATCTTTAACTCCTCTGTGTAGTTGTCTCTCCAAATTTGAAGCTCTAAACTA 170763	
Qy	61	CAAAATTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGACTCCACCGTGGACCC 120	
Db	170762	CAAAATTGTTCTTCAATGGAGCCCCAGACGACGCTCCATGACTAAGACTACCGCAGACC 170703	
Qy	121	CTGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 180	
Db	170702	CTGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTCGAAGGCACCCCTCCAGAG 170643	
Qy	181	GAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240	
Db	170642	GAATCTCAACTGCATACCCCACTACGCCCAATTCACAGAGAGCAGTTAGACAGT 170583	
Qy	241	CATCAGCCAACTCCCCAAGCAGCACTTGGGTTTTCTGTTGAGAGGGGGACTTGAGAGAC 300	
Db	170582	CATCGGCATCTCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGTACTCAGAGAC 170523	
Qy	301	AGGACTAGCTGGATTTCTTAGGCCCAACAGAGATCCCTAGCCCTAGCTGGAGAGTGACT 360	
Db	170522	AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGACCTAGCTGGAGAGTGACT 170463	
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCGAGCAATC----- 411	
Db	170462	GCTTCTACCTTTTAAACCCGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGGTAGAA 170403	
Qy	412	AGAGAGCTCACTAAAAATGCTAATTAGGCAAAAAATAGAGGTAAGAAATAGCCAATCATC 471	
Db	170402	AGAGAGCTCACTAAAAATGCTAAGCTAGGCTAAAAACAGAGGTAAGAAATAGCCAATCATC 170343	
Qy	472	TATTGCTTGAGACACAGCGGGAGGACAGGATCGGGATATAAACCCAGGCATTGCGAGC 531	
Db	170342	TATCGCCTGAGAACACACAGTGGGAGGACAAATGATGGGTTTATTAACCCAGGCATTGCGAGC 170283	
Qy	532	CGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTTCACTCTA 591	
Db	170282	CAGCAATGGCTACCATTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCACTCTA 170223	
Qy	592	TTTCACTCTATTAAATCTTGCACCTGAAAAAAGGAAAAA 635	
Db	170222	TTTCACTCTATTAAATCTTGCACAGCAAAAAAAGAAAAA 170179	
RESULT 17			
ADQ94981/c			
ID	ADQ94981	standard; DNA; 326014 BP.	
AC	ADQ94981;		
XX	23-SEP-2004	(first entry)	
DT	Human kinase genomic DNA.		
DE	Kinase; diagnosis; cancer; tissue growth abnormality;		
KW	drug screening assay; cytostatic; gene therapy; human; SNP;		
KW	single nucleotide polymorphism; gene; ds.		
XX	Homo sapiens.		
OS	Location/Qualifiers		
XX	Key	replace(1138..1140,GT)	
FT	variation	/*tag= a	
FT		/standard_name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(1399,G)	
FT		/*tag= b	
FT	variation	replace(1871..1873,TG)	
FT		/*tag= c	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"	
FT		replace(1875..1877,TA)	
FT	variation	/*tag= d	
FT		/standard_name= "Single nucleotide polymorphism (SNP)"	
FT	CDS	3000..323016	
FT		/*tag= f	
FT		/product= "Human kinase protein"	
FT		/transl_except= (pos:3024..3026, aa:Arg)	
FT		/transl_except= (pos:142261..142263, aa:Glu)	
FT		/transl_except= (pos:142261..144102, aa:Glu-Ser)	
FT	exon	3000..3169	
FT		/*tag= e	
FT	intron	3170..53737	
FT		/*tag= g	
FT	variation	replace(3180,A)	
FT		/*tag= h	
FT		/standard_name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(7590,A)	
FT		/*tag= i	
FT		/standard_name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(9315,C)	
FT		/*tag= j	
FT		/standard_name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(9572,A)	
FT		/*tag= k	
FT		/standard_name= "Single nucleotide polymorphism (SNP)"	

```
FT variation replace(11004,G)
FT /tag= l
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(11378..11380,AA)
FT /tag= m
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(12478,A)
FT /tag= n
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(12485..12487,TC)
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(13904,G)
FT /tag= p
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(12444..12446,TT)
FT /tag= q
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(26805,A)
FT /tag= r
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28010,R)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28107,Y)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28178,R)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28356,T)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(29535,G)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(31808,T)
FT /tag= x
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(32807,C)
FT /tag= y
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(34531,G)
FT /tag= z
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(35318,G)
FT /tag= aa
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(35463,G)
FT /tag= ab
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(35477,T)
FT /tag= ac
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37050,G)
FT /tag= ad
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37470,G)
FT /tag= ae
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37481,A)
FT /tag= af
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37499,T)
FT /tag= ag
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(40177..40179,AA)
FT /tag= ah
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(40178,A)
FT /tag= ai
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(40787,A)
FT /tag= aj
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(44582,G)
FT /tag= ak
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(45615,A)
FT /tag= al
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(46091,T)
FT /tag= am
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(47844,C)
FT /tag= an
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(48013,T)
FT /tag= ao
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(48072,A)
FT /tag= ap
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(50802,T)
FT /tag= aq
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(53305,A)
FT /tag= ar
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(53378..53380,GG)
FT /tag= as
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 53738..53903
FT /tag= at
FT 53904..83114
FT /tag= au
FT replace(56233,G)
FT /tag= av
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(56453,C)
FT /tag= aw
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(59300,T)
FT /tag= ax
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(60998,A)
FT /tag= ay
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(63013,A)
FT /tag= az
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(63293,A)
FT /tag= ba
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(64181,T)
FT /tag= bb
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(64191,A)
FT /tag= bc
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(68992,C)
FT /tag= bd
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(76246,Y)
FT /tag= be
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(7368,G)
FT /tag= bf
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(78043,A)
FT /tag= bg
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(78395,T)
FT /tag= bh
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(79474,G)
FT /tag= bi
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation
```


QY 481 AGGACACAGGGGAGGACAGGATCGGATATACCCAGGCAATTCGAGCCGGCAACG 540
 Db 2625 AGAGCACAGGAGGAGGACAAATGATCGGATATACCCAGGCAATTCGAGCCGGCAACG 2684
 QY 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTTATTTCACTCT 600
 Db 2685 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGAGCTCTGTTTCACTCTTATTTCACTCT 2744
 QY 601 ATTAATCTTGAACCTGAAAAAATAAAAAA 635
 Db 2745 ATTAATCTTGAACCTGAAAAAATAAAAAA 2779

RESULT 19

AA25660
 ID AAX25660 standard; cDNA to mRNA; 1136 BP.

AC AAX25660;

DT 21-MAY-1999 (first entry)

DE Human endogenous retrovirus W clone cl.C4C5.

XX Clone; human endogenous retrovirus; genome; autoimmune disease;
 KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
 KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

XX Human endogenous retrovirus.

XX WO9902696-A1.

XX 21-JAN-1999.

XX 06-JUL-1998; 98WO-FR001442.

XX 07-JUL-1997; 97FR-00008815.

XX (INMR) BIO MERIEUX.

XX Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;

XX WPI; 1999-120897/10.

XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
 PT disease, and abnormal or failed pregnancy.

XX Claim 1; Page 59-60; 106pp; French.

XX This sequence represents clone cl.C4C5 of the human endogenous retrovirus
 CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
 CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
 CC dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility

XX Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;

Query Match 82.8%; Score 525.8; DB 2; Length 1136;
 Best Local Similarity 90.7%; Fred. No. 1.3e-152;
 Matches 576; Conservative 0; Mismatches 47; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTAACTCTTGTGTTAAGTTTCTCTTCCAGAAATCAAAACTGTAAACTA 60
 Db 510 CCCTGTATCTTAACTCTTGTGTTAAGTTTCTCTTCCAGAAATCGAGCTGTAAACTA 569
 QY 61 CAAATTTGTTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
 Db 570 -----CAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGGAGACCC 617

QY 121 CTGGACCGGCTCTGTAGCCCATGCTCCGATGTAAATGACATTTGAAGGCAACCTCTCCGAG 180
 Db 618 CTGGACCGGCTCTGTAGCCCATGCTCCGATGTAAATGACATTTGAAGGCAACCTCTCCGAG 677
 QY 181 GAAATCTCACTGACACACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 240
 Db 678 GAAATCTCACTGACACACCTCTACTACGCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 737
 QY 241 CATCAGCCAACTCTCCCAACAGCAGCTTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGAC 300
 Db 738 CGTGGCCAACTCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGAGCTGAGAGAC 797
 QY 301 AGGACTAGCTGGATTCTTCAGCCAAAGAAATCCCTAAGCCTAGCTAGCGAAGGTGACT 360
 Db 798 AGGACTAGCTGGATTCTTCAGCTGACTAAGAAATCCCTAAGCCTAGCTAGCGAAGGTGACC 857
 QY 361 GCATCCACCTCTTAAACATGGGCTTGCACCTTAGCTCAGACCCGACCAATCAGAGAGCTC 420
 Db 858 ACATCCACCTTTAAACACAGCGGCTTGCACCTTAGCTCAGACCTGACCAATCAGAGAGCTC 917
 QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTG 480
 Db 918 ACTAAATGCTAATTAGGCAAGAGGTAAGAAATAGCAATCATCTATTGCTG 977
 QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGATATATAACCCAGGCAATTCGAGCCGGCAACG 540
 Db 978 AGAGCACAGCGGAGGAGCAATGATCGGATATATAACCCCAAGTCTTTTCGAGCCGGCAACG 1037
 QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTTATTTCACTCT 600
 Db 1038 CAACCCCTTTGGTCCCTCCCTTTGATGGGAGCTCTGTTTCACTCTTATTTCACTCT 1097
 QY 601 ATTAATCTTGAACCTGAAAAAATAAAAAA 635
 Db 1098 ATTAATCTTGAACCTGCGAAAAAATAAAAAA 1132

RESULT 20

AA59210
 ID AA59210 standard; DNA; 1136 BP.

XX AC AA59210;

XX 07-NOV-2000 (first entry)

XX 3' pol gene and 3' non coding sequences of HERV-W from human genome.

XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
 KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX Homo sapiens.

XX WO200043521-A2.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-FR000144.

XX 21-JAN-1999; 99FR-00000888.

XX (INMR) BIO MERIEUX.

XX Paranhos-Baccala G, Mallet F, Voisset C;

XX WPI; 2000-499229/44.

XX New nucleic acid from human endogenous retrovirus, useful e.g. for
 .PT diagnosis of autoimmune disease and complications of pregnancy, contains
 PT at least part of the gag gene.

XX Disclosure; Page 46; 53pp; French.

XX The present sequence represents an endogenous retroviral nucleic acid

CC fragment, which is associated with an autoimmune disease, and is
CC integrated into the human genome. The fragment is originally derived from
CC a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
CC retrovirus is associated with autoimmune disease, failure of pregnancy or
CC disorders of pregnancy. The nucleic acid fragment, or proteins derived
CC from it, are useful for diagnosis of autoimmune disease (specifically
CC multiple sclerosis) and for monitoring pregnancy. The nucleic acid
CC fragments may also be used for in situ labelling of isolated chromosomes,
CC while the transcription product can be used to study or monitor T cell
CC proliferation in vitro
XX
SQ Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;

Query Match 82.8%; Score 525.8; DB 3; Length 1136;
Best Local Similarity 90.7%; Pred. No. 1.3e-152;
Matches 576; Conservative 0; Mismatches 47; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60
DB |||||
DB 510 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 569
QY 61 CAAATTTGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB |||||
DB 570 -----CAAAATGGAGCCCAAGATGCATCCAGACTAAGATCTACCGCAGACCC 617
QY 121 CTGGACCGGCTCTAGCCATGCTCCGATGTTTAATGACATGAAGGACCCCTCCCGAG 180
DB |||||
DB 618 CTGGACCGGCTCTAGCCATGCTCCGATGTTTAATGACATCAAAAGGACCCCTCCGAG 677
QY 181 GAAATCTCAACTGCACAACTCTACTATGCTCCCAATTCAGCGGAGCAGTTAGACGGT 240
DB |||||
DB 678 GAAATCTCACTGCACAACTCTACTATGCTCCCAATTCAGCGGAGCAGTTAGACGGT 737
QY 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTCTGTGAGAGGGGAGCTGAGAGAC 300
DB |||||
DB 738 CGTGGCCCACTCCCAACAGACACTTAGGTTTCTCTGTGAGATGGGGAGCTGAGAGAC 797
QY 301 AGGACTAGCTGGATTTCTTAGGCCCAAGAAATCCCTAAGCTAGCTGGAAGTGACT 360
DB |||||
DB 798 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACC 857
QY 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
DB |||||
DB 858 ACATCCACCTTTAAACAGGGGGCTTGCACTTAGTTTCAACCTGACCAATCAGAGAGCTC 917
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
DB |||||
DB 918 ACTAAATGCTAATTAGGCAAAAGACAGAGGTAAAGAAATAGCCAAATCATCTATTGCATG 977
QY 481 AGAGCAGCGGGAGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAACGG 540
DB |||||
DB 978 AGAGCAGCAGGAGGAGCAATGATCGGATATAAACCAGGCTCTTCGAGCGGCAACGG 1037
QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGGCTCTGTTTTCATCTCTATTTCATCTCT 600
DB |||||
DB 1038 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCATCTCT 1097
QY 601 ATTAATTTGCAACTGAAAAAATAAAAAAAAAAAAAA 635
DB |||||
DB 1098 ATTAATTTGTCAGCTCGAAAAAATAAAAAAAAAAAAAA 1132

RESULT 21
AA25661
ID AA25661 standard; cDNA to mRNA; 2782 BP.
XX AC
XX AA25661;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.PH74.
XX
XX Clone; human endogenous retrovirus; genome; autoimmune disease;
KW

KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.
XX
PP 06-JUL-1998; 98WO-FR001442.
XX
PR 07-JUL-1997; 97FR-00008815.
XX
XX (INMR) BIO MERIEUX.
XX
PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX
XX WPI; 1999-120897/10.
XX
XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
XX exclusively in placenta and useful in diagnosis and therapy of autoimmune
XX disease, and abnormal or failed pregnancy.
XX
XX Claim 1; Page 60-63; 106pp; French.
XX
XX This sequence represents clone cl.PH74 of the human endogenous retrovirus
XX (HERV) W genome. The nucleic acids, their fragments or peptides encoded
XX by them are markers of autoimmune disease (e.g. multiple sclerosis,
XX rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
XX dependent diabetes and related pathologies) and of abnormal or
XX unsuccessful pregnancy and can be used as chromosomal markers for
XX susceptibility to these conditions, or proximity markers of genes
XX associated with this susceptibility
XX
SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 82.3%; Score 522.6; DB 2; Length 2782;
Best Local Similarity 90.4%; Pred. No. 2e-151;
Matches 574; Conservative 0; Mismatches 49; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60
DB |||||
DB 2158 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 2217
QY 61 CAAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
DB |||||
DB 2218 -----CAAAATGGAGCCCAAGATGCAGTCCAGACTAAGATCTACCGCAGACCC 2265
QY 121 CTGGACCGGCTCTAGCCATGCTCCGATGTTTAATGACATGAAGGACCCCTCCCGAG 180
DB |||||
DB 2266 CTGGACCGGCTCTAGCCATGCTCCGATGTTTAATGACATGAAGGACCCCTCCCGAG 2325
QY 181 GAAATCTCAACTGCACAACTCTAGTCCCAATTCAGCGGAGCAGATTAGAGCGGT 240
DB |||||
DB 2326 GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGCAGGAGCAGATTAGAGCGGT 2385
QY 241 CATCAGCCAACTCTCCCAACAGACACTTGGGTTTCTCTGTGAGAGGGGAGCTGAGAGAC 300
DB |||||
DB 2386 GGTCCGCCAACCTCTCCCAACAGCAGCTTAGTTTCTCTGTGAGATGGGGAGCTGAGAGAC 2445
QY 301 AGGACTAGCTGGATTTCTTAGCCCAACGAGATCCCTAAGCTAGCTGGGAGGTGACT 360
DB |||||
DB 2446 AGGACTAGCTGGATTTCTTAGCTGACTAAGAAATCTTAAAGCTAGGTGGGAGGTGACC 2505
QY 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
DB |||||
DB 2506 ACATCCACCTTTAAACACGGGCTTGCACTTAGCTTCACCTGACCAATCAGAGAGCTC 2565
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
DB |||||
DB 2566 ACTAAATGCTAATTAGGCAAAAGAGGAGGTAAAGAAATAGCCAAATCATCTATTTCCTG 2625
QY 481 AGAGCAGCGGGAGGAGCAGAGTCCGGATATAAACCAGGCAATTCGAGCGGCAACGG 540

CC choriocarcinoma, hydatiform mole, placental site tumour and missed/
CC incomplete abortion). Syncytin is a human gene derived from the envelope
CC gene of human endogenous defective retrovirus, HERV-W. The present
CC invention is based partly on the discovery that syncytin expression is
CC dramatically reduced in preclampsia, and is also mis-localised to the
CC apical syncytiotrophoblast membrane. The present sequence is human
CC syncytin cDNA

XX Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;

Query Match 81.5%; Score 517.4; DB 6; Length 2930;
Best Local Similarity 91.4%; Pred. No. 8.3e-150;
Matches 564; Conservative 0; Mismatches 41; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTTCCAGAAATCAAAATCTGTAAACTA 60
DB |||||
2325 CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTTCCAGAAATCGAAGCTGTAAACTA 2384
DB |||||
QY 61 CAAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB |||||
2385 -----CAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 2432
QY 121 CTGACCGGCTGTAGCCCATGCTCGATGTTAAATGACATTAAGAGCACCCTCCCGAG 180
DB |||||
2433 CTGACCGGCTGTAGCCCATGCTCGATGTTAAATGACATTAAGAGCACCCTCCCGAG 2492
QY 181 GAATCTCACTCAGACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGT 240
DB |||||
2493 GAAATCTCAGCTCAGACACCTCTACTACGCCCAATTCAGCGGGAAGCAGTTAGAGCGT 2552
QY 241 CATCAGCAACCTCCCCCAACAGCACCTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB |||||
2553 CGTCGGCAACCTCCCCCAACAGCACCTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 2612
QY 301 AGCACTAGCTGGATTTCTTAGGCAACGAAAGATCCCTAAGCCCTAGCTGGGAAGTGACT 360
DB |||||
2613 AGCACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCCTAGCTGGGAAGTGACT 2672
QY 361 GCATCCACCTTAAACATGGGCTTGCATCTAGCTCACACCCGACCAATCAGAGAGCTC 420
DB |||||
2673 ACATCCACCTTAAACATGGGCTTGCATCTAGCTCACACCCGACCAATCAGAGAGCTC 2732
QY 421 ACTAAATGCTAATTTAGGCAAAATAGAGTAAAGAAATAGCAATCATCTATTGCGCTG 480
DB |||||
2733 ACTAAATGCTAATTTAGGCAAAACAGAGGTAAGAAATAGCAATCATCTATTGCGCTG 2792
QY 481 AGACACAGCGGGAGGACAAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGCAACGG 540
DB |||||
2793 AGACACAGCGGGAGGACAAATGATCGGGATATAAACCCAGGCAATTCGAGCCGCAACGG 2852
QY 541 CAACCCCTTTGGTCCCTCCCTCTTGTATGGGCGCTCTTTTCACTCTATTTCACCTCT 600
DB |||||
2853 CAACCCCTTTGGTCCCTCCCTCTTGTATGGGCGCTCTTTTCACTCTATTTCACCTCT 2912
QY 601 ATTAATCTTGCAACTG 617
DB |||||
2913 ATTAATCTTGCAACTG 2929

RESULT 25
ABL61744
ID ABL61744 standard; DNA; 56093 BP.
XX ABL61744;
AC ABL61744;
DT 15-MAY-2002 (first entry)
XX Colon adenocarcinoma related gene sequence SEQ ID NO:81.
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.
OS WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0231133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 22-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236112P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 81; 44pp; English.
PS The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, adenocarcinoma, papillary carcinoma and Wilm's
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;

Query Match 81.3%; Score 516.2; DB 6; Length 56093;
Best Local Similarity 91.1%; Pred. No. 8.3e-149;
Matches 564; Conservative 0; Mismatches 43; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTTAACTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60
DB 37274 CCTGTATCTTTAACTCTCTTAACTTGTCTCTCCAGAAATCGAAGCTGTAAACTA 37333
QY 61 CAAATTTGTTTCAAAATGAGCACCAGATGAGTCCATGACTAAGATCCACCGTGACCC 120
DB 37334 -----CAAATGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGACCC 37381
QY 121 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGAGGACCCCTCCGAG 180
DB 37382 CTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAAAAGGACCCCTCCGAG 37441
QY 181 GAAATCTCAACTGCACAAACCTTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
DB 37442 GAAATCTCAAGTGCACAACTTACTACGCCCAATTCAGCAGGAGCAGTTAGAGCGGT 37501
QY 241 CATCAGCCAACTCCCAACAGCATTGGGTTTCTGTGAGAGGGGGGACTGAGAGAC 300
DB 37502 CGTCGGCCAACTCCCAACAGCATTAGGTTTCTGTGAGATGGGGGACTGAGAGAC 37561
QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAGATCCCTAAGCTAGCTGGGAGGTGACT 360
DB 37562 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCTAGCTGGGAGGTGACC 37621
QY 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCGACCAATTCAGAGAGCTC 420
DB 37622 ACATCCACCTTTAAACAGGGGCTTGCACTTAGCTCACACCTGACCAATTCAGAGAGCTC 37681
QY 421 ACTAAATGCTAATTAGGCAAAAATAGAGGTAAAGAAATAGCCAATCATCTATTGCTCG 480
DB 37682 ACTAAATGCTAATTAGGCAAAAGAGGAGTAAAGAAATAGCCAATCATCTATTGCTCG 37741
QY 481 AGAGCAGCGGGGAGGACAGGATCGGATATAAACCCAGGCATTCGAGCCCGCAACGG 540
DB 37742 AGAGCAGCAGGAGGAGGACAAATGATCGGGATATAAACCCAAAGTCTTCGAGCCGGCAACGG 37801
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 37802 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 37861
QY 601 ATTAATCTTGCAACTGAA 619
DB 37862 ATTAATCTTGCAACTGCA 37880

RESULT 26
AAAX25665
ID AAAX25665 standard; cDNA to mRNA; 7582 BP.
XX
AC AAAX25665;
XX
DT 21-MAY-1999 (first entry)
XX Complete human endogenous retrovirus W genome.
DE
XX Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW

KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX Human endogenous retrovirus.
OS
XX WO9902696-A1.
PN
XX 21-JAN-1999.
PD
XX
XX 06-JUL-1998; 98WO-FR001442.
PF
XX 07-JUL-1997; 97FR-00008815.
PR
XX (INMR) BIO MERIEUX.
PA
XX Beesme F, Blond J, Bouton O, Mandrand B, Mallet P;
FI
XX WPI; 1999-120897/10.
DR
XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
XX exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.
PT
XX
PS
XX Claim 1; Page 71-74; 106pp; French.

XX This sequence represents the complete sequence of the human endogenous
CC retrovirus (HERV) W genome. The nucleic acids, their fragments or
CC peptides encoded by them are markers of autoimmune disease (e.g. multiple
CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
CC insulin-dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility

XX Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;
SQ
Query Match 80.5%; Score 511.4; DB 2; Length 7582;
Best Local Similarity 89.5%; Pred. No. 9.8e-148;
Matches 552; Conservative 11; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTTAACTTGTCTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 6976 CCTGTATCTTTAACTCTCTTAACTTGTCTCTTCCAGAAATCGAAGCTGTAAACTA 7035
QY 61 CAAATTTGTTCTTAAATGGAGCACCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 7036 -----CAAATGGAGCCCAAGATGAGTCCAGAGCTAAGATCTACCGCAGACCC 7083
QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCACCCCTCCCGAG 180
DB 7084 CTGGACCGGCTGCTAGCCCATGCTGATGTTAATGACATCAAGGCACCCCTCCTGAG 7143
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 7144 GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
QY 241 CATCAGCCAACTCCCAACAGCATTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
DB 7204 SGTGCGCAACCTCCCAACAGCATTAGGTTTCTGTTGAGATGGGGGACTGAGAGAC 7263
QY 301 AGGACTAGCTGGATTTCTAGCCCAACGAAATCCCTAAGCTAGCTGGGAAGGTGACT 360
DB 7264 AGGACTAGCTGGATTTCTAGCTGATTAAGAAATCCYTAAGCTAGTGGGAAGGTGACT 7323
QY 361 GCATCCACCTCTAAACATGGGCTTTGCAACTTTAGCTCAACCCGACCAATTCAGAGAGCTC 420
DB 7324 ACATCCACCTTTAAACACGGGGCTTGCACTTAGTCTACACCTGACCAATTCAGAGAGCTC 7383
QY 421 ACTAAATGCTAATTAGGCAAAAATAGAGGTAAAGAAATAGCCAATCATCTATTGCTCG 480
DB 7384 ACTAAATGCTAATTAGGCAAAAGAGCAGGAGGTAAAGAAATAGCCAATCATYATTGCMFG 7443
QY 481 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGCATTCAGAGCCGCAACGG 540

Db 7444 AGAGCAGCAGGAGGACAATGATCGGATATATAACCAAGTTCGAGCGGCAACGG 7503
Qy 541 CAACCCCTTTGGTCCCTCCCTTTGATAGGCGCTCTGTTTCACTCTATTTCACTCT 600
Db 7504 CAACCCCTTTGGTCCCTCCCTTTGATAGGAGCTCTGTTTCACTCTATTTCACTCT 7563
Qy 601 ATTAATCTTGCACACTG 617
Db 7564 ATTAATCTTGCACACTG 7580

RESULT 27
ID AAA59215 standard; DNA; 7582 BP.
XX AC AAA59215;
XX 07-NOV-2000 (first entry)
XX Human endogenous retrovirus W (HERV-W) sequence.
XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX Human endogenous retrovirus.
XX Key Location/Qualifiers
FH LTR 1. .120
FT /*tag= a
FT /note= "R of 5' LTR"
FT 121..575
FT /*tag= b
FT /note= "US of 5' LTR"
FT primer_bind 579..596
FT /*tag= c
FT CDS 5581..7194
FT /*tag= d
FT /note= "ORF1 env538"
FT 7039..7194
FT /*tag= e
FT CDS 7112..7255
FT /note= "ORF2 52 AA"
FT /*tag= f
FT /note= "ORF3 48 AA"
FT misc_feature 7244..7254
FT /*tag= g
FT /note= "polypurine tract"
FT LTR 7256..7582
FT /*tag= h
FT polyA_signal /note= "U3-R of 3' LTR"
FT 7563..7569
FT /*tag= i

XX WO200043521-A2.
XX 27-JUL-2000.
XX 21-JAN-2000; 2000WO-FR000144.
XX 21-JAN-1999; 99FR-00000888.
XX (INNR) BIO MERIEUX.
XX Paranhos-Baccala G, Mallet F, Voisset C;
XX WPI; 2000-499229/44.
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy, contains
PT at least part of the gag gene.
XX PS Disclosure; Page 49-52; 53pp; French.

CC The present sequence represents an endogenous retrovirus, which is
CC associated with an autoimmune disease, and is integrated into the human
CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
CC proteins derived from it, are useful for diagnosis of autoimmune disease
CC (specifically multiple sclerosis) and for monitoring pregnancy. The
CC nucleic acid fragments may also be used for in situ labelling of isolated
CC chromosomes, while the transcription product can be used to study or
CC monitor T cell proliferation in vitro
XX
SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;

Query Match 80.5%; Score 511.4; DB 3; Length 7582;
Best Local Similarity 89.5%; Pred. No. 9.8e-148;
Matches 552; Conservative 11; Mismatches 42; Indels 12; Gaps 1;
Qy 1 CCTGTATCTTTAACTCCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
Db 6976 CCTGTATCTTTAACTCCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 7035
Qy 61 CAAATTGTTCTTCAATGGAGCAGACGATGGAGTCCATGACTAAGATCCACCCGTGACCC 120
Db 7036 -----CAAATGGAGCCCAAGATGCAGTCCAGACTTAAGATCTACCCGACCC 7083
Qy 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGGCAACCCCTCCCGAG 180
Db 7084 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGGCAACCCCTCCCGAG 7143
Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 7144 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 7203
Qy 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 7204 SGTCCGCCCAACCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 7263
Qy 301 AGGACTAGCTGGATTTCTAGGCCAACGAAGATCCCTAAGCCTAGCTGGAAAGGTGACT 360
Db 7264 AGGACTAGCTGGATTTCTAGGCCAACGAAGATCCCTAAGCCTAGCTGGAAAGGTGACT 7323
Qy 361 GCATCACCCTCTAAACATGGGGCTTGCACTTAGCTCAGCCGACCAATCAGAGAGCTC 420
Db 7324 ACATCCACCCTTTAAACACGGGGCTTGCACTTAGCTCAGCCGACCAATCAGAGAGCTC 7383
Qy 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
Db 7384 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 7443
Qy 481 AGAGCAGCGGAGGAGCAAGGATCGGATATAAAACCCAGGCAATTCAGCCCGCAACGG 540
Db 7444 AGAGCAGCGGAGGAGCAAGGATCGGATATAAAACCCAGGCAATTCAGCCCGCAACGG 7503
Qy 541 CAACCCCTTTGGTCCCTCCCTTTGATAGGCGCTCTGTTTCACTCTATTTCACTCT 600
Db 7504 CAACCCCTTTGGTCCCTCCCTTTGATAGGAGCTCTGTTTCACTCTATTTCACTCT 7563
Qy 601 ATTAATCTTGCACACTG 617
Db 7564 ATTAATCTTGCACACTG 7580

RESULT 28
ABN97978
ID ABN97978 standard; DNA; 46340 BP.
XX AC ABN97978;
XX AC ABN97978;
XX 01-AUG-2002 (first entry)
XX Human retroviral sequence H13.
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;

413 RGGGCTCACTAAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATC 471

QY 61 CAAATTTGTTCTTCAAAATGGAGCACCAGATGGAGTCCAGTCAAGATCCACCGTGACCC 120
 DB 100551 CAAATCGTTCTTCAAAATGGAGCACCAGATGAAGTCCATGACTAAGATCTACCGTGACCC 100492
 QY 121 CTGGACCGGCTCTAGCCCATCTCCGATGTTAAATGACATTTGAAGGACCCCTCCCGAG 180
 DB 100491 CTGGACCGGCTCTAGCCCATCTCCGATGTTAAATGACATTTGAAGGACCCCTCCCGAG 100432
 QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGACGGT 240
 DB 100431 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGACGGT 100372
 QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
 DB 100371 CGTCAGCAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 100312
 QY 301 AGGACTAGCTGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 DB 100311 AGGACTAGCTGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 100252
 QY 361 GATCACCCTCTAAACATGGGGTTGCAACTTAGCTCAGCCCGACCAATC----- 411
 DB 100251 ACATCCACCTTTAAACATGGGGTTGCAACTTAGCTCAGCCCGACCAATCAGGTAGTAA 100192
 QY 412 AGAGAGCTCACTAAATGCTAATTAGCAAAATAGAGGTAAAGAAATAGCCCAATCATC 471
 DB 100191 AGAGAGCTCACTAAATGCTAATTAGCAAAATAGAGGTAAAGAAATAGCCCAATCATC 100132
 QY 472 TATTGCTGAGAGCAGACGCGGAGGACAAAGGATCGGATATAAACCCAGGCAATTCGAGC 531
 DB 100131 TATCGCTGAGAGCAGACGCGGAGGACAAATGATCGGATATAAACCCAGGCAATTCGAGC 100072
 QY 532 CGGCAACGGCAACCCCTTTGGGTCCCTCTTGTATGGGCGCTCTGTTTCACTCTA 591
 DB 100071 CGGCAACGGCTACCTTTTGGGTCCCTCTTGTATGGGAGCTCTCT-----GT 100018
 QY 592 TTTCACTCTAATAATCTTCAACTGAAAAAATAATA 630
 DB 100017 CTTCACTCTAATAATCTTCAACTGAAAAAATAATA 99979

RESULT 30

ABN97929 standard; DNA; 10499 BP.

AC ABN97929;

01-AUG-2002 (first entry)

Human retroviral sequence HERV-7q.

Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;

multiple sclerosis; ds.

Human endogenous retrovirus.

WO9967395-A1.

29-DEC-1999.

23-JUN-1999; 99WO-FR001513.

23-JUN-1998; 98FR-00007920.

(INRM) INSERM INST NAT SANTE & RECH MEDICALE.

Alliel PM, Perin J, Rieger F;

WPI; 2000-160587/14.

New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used

for diagnosis, treatment and prevention of autoimmune and neurological

PT

XX diseases.

XX Claim 3; Fig 1; 225bp; French.

XX The present invention relates to new nucleic acid sequences of human
 CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
 CC Regulatory elements associated with HERV-7q may alter expression of other
 CC genes (even remote genes) on the same chromosome, inducing immunological
 CC and/or neurological changes (which may be pathological or protective/
 CC curative). HERV-7q peptides can be used to improve efficiency of the
 CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
 CC sequences can be used in immunogenic or vaccinating compositions, for
 CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention

XX Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 U; 0 Other;

XX Query Match 79.7%; Score 505.8; DB 3; Length 10499;

XX Best Local Similarity 91.1%; Pred. No. 6.3e-146;

XX Matches 564; Conservative 0; Mismatches 42; Indels 13; Gaps 2;

QY 1 CCTGTATCTTTAACTCCTCTGTTAAGTTTGTCTCTCCAGAAATCAAAAATCTGTAAACTA 60

DB 9274 CCTGTATCTTTAACTCCTCTGTTAAGTTTGTCTCTCCAGAAATCAAAAATCTGTAAACTA 9333

QY 61 CAAATTTGTTCTTCAAAATGGAGCACCAGATGGAGTCCAGTCAAGATCCACCGTGACCC 120

DB 9334 -----CAAATGGAGCACCAGATGGAGTCCAGTCAAGATCCACCGTGACCC 9381

QY 121 CTGGACCGGCTCTAGCCCATCTCCGATGTTAAATGACATTTGAAGGACCCCTCCCGAG 180

DB 9382 CTGGACCGGCTCTAGCCCATCTCCGATGTTAAATGACATTTGAAGGACCCCTCCCGAG 9441

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGACGGT 240

DB 9442 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGACGGT 9501

QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300

DB 9502 C-TCGGCCAAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 9560

QY 301 AGGACTAGCTGGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

DB 9561 AGGACTAGCTGGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 9620

QY 361 GCATCCACCTCTAAACATCGGGCTTGCAACTTAGCTCAGACCCGACCAATCAGAGAGCTC 420

DB 9621 ACATCCACCTTTAAACATCGGGCTTGCAACTTAGCTCAGACCCGACCAATCAGAGAGCTC 9680

QY 421 ACTAAATGCTAATAGGCAAAAATAGGAGTTAAGAAATAGCAATCATCTATTGCTG 480

DB 9681 ACTAAATGCTAATAGGCAAAAATAGGAGTTAAGAAATAGCAATCATCTATTGCTG 9740

QY 481 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCCCGCAACGG 540

DB 9741 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCCCGCAACGG 9800

QY 541 CAAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTTCACTCTATTCTCTCT 600

DB 9801 CAAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTTCACTCTATTCTCTCT 9860

QY 601 ATTAAATCTTGAACCTGAA 619

DB 9861 ATTAAATCTTGAACCTGCA 9879

RESULT 31

ACN44958

ID ACN44958 standard; DNA; 285020 BP.

XX

[illegible]

SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 Query Match 78.7%; Score 499.6; DB 4; Length 1894;
 Best Local Similarity 88.7%; Pred. No. 2.3e-144;
 Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 60
 DB 1217 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 1276
 QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGAGTCAATGATCAAGATCCACCTGAGACC 120
 DB 1277 CAAATGTTCTTCAAAATGAGACACAGATGAGTCAATGATCAAGATCCACCTGAGACC 1336
 QY 121 CTGACCGGCTCTGATCCCATGCTCGATGTTAAATGACATTCGAAGCACCCTCCCGAG 180
 DB 1337 CTGACCGGCTCTGATCCCATGCTCGATGTTAAATGACATTCGAAGCACCCTCCCGAG 1396
 QY 181 GAAATCTCAACTGCACAACTCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 DB 1397 GAAATCTCAACTGCACAACTCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGTGT 1456
 QY 241 CATGAGCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
 DB 1457 TGTGGCCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 1516
 QY 301 AGGACTAGCTGGATTTCTAGGCGCAGAGAAATCCCTTAGCCCTAGCTGGGAAGGTGACT 360
 DB 1517 AGGAATAAATAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGGTGACC 1576
 QY 361 GCATCCACTCTAAACATGCGGCTTGCAACTTACTAGCTCACACCGGACCAATC----- 411
 DB 1577 GCTTCCACTTTAAACACCGGCTTGCAACTTACTAGCTCACACCGGACCAATCAGNACTAA 1636
 QY 412 AGAGAGCTCAATAATGCTAAATAGGCAAAATAGAGGTAAAGAAATAGCCAAATCATC 471
 DB 1637 AGAGAGCTCAATAATGCTAAATAGGCAAAATAGAGGTAAAGAAATAGCCAAATCATC 1696
 QY 472 TATTGCTGAGACAGACGCGGAGGACAGGATCGGGATATAACCCAGGCATTTCGAGC 531
 DB 1697 TGTGCTGACAGACAGCAGGAGGAGCAATGATCGGGATATAACCCAGGCATTTCGAGC 1756
 QY 532 CGGCAACGCGCAACCTCTTGGGTCCCTCTTGTATGGGCGCTCTGTTTTCACCTCTA 591
 DB 1757 CAGTACAGTACCTCTTGGGTCCCTCTTGTATGGGAGCTCTGT----- 1806
 QY 592 TTTCACTCTATTAAATCTTGCAACTGCAAAATAAATAAGAAA 635
 DB 1807 CTTCACTCTATTAAATCTTGCAACTGCAAAATAAATAAGAAA 1850

RESULT 33
 ABA56337
 ID ABA56337 standard; DNA; 1894 BP.
 XX
 AC ABA56337;
 XX
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #4642.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 FN WO200157277-A2.
 XX
 XX
 PD 09-AUG-2001.
 XX
 XX
 PF 30-JAN-2001; 2001WO-US000669.
 XX
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR

PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX WPI; 2001-483447/52.
 DR
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human foetal liver.
 PS
 XX Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human foetal liver. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 Query Match 78.7%; Score 499.6; DB 4; Length 1894;
 Best Local Similarity 88.7%; Pred. No. 2.3e-144;
 Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 60
 DB 1217 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 1276
 QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGAGTCAATGATCAAGATCCACCTGAGACC 120
 DB 1277 CAAATGTTCTTCAAAATGAGACACAGATGAGTCAATGATCAAGATCCACCTGAGACC 1336
 QY 121 CTGACCGGCTCTGATCCCATGCTCGATGTTAAATGACATTCGAAGCACCCTCCCGAG 180
 DB 1337 CTGACCGGCTCTGATCCCATGCTCGATGTTAAATGACATTCGAAGCACCCTCCCGAG 1396
 QY 181 GAAATCTCAACTGCACAACTCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 DB 1397 GAAATCTCAACTGCACAACTCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGTGT 1456
 QY 241 CATGAGCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
 DB 1457 TGTGGCCAACTCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 1516
 QY 301 AGGACTAGCTGGATTTCTAGGCGCAGAGAAATCCCTTAGCCCTAGCTGGGAAGGTGACT 360
 DB 1517 AGGAATAAATAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGGTGACC 1576
 QY 361 GCATCCACTCTAAACATGCGGCTTGCAACTTACTAGCTCACACCGGACCAATC----- 411
 DB 1577 GCTTCCACTTTAAACACCGGCTTGCAACTTACTAGCTCACACCGGACCAATCAGNACTAA 1636
 QY 412 AGAGAGCTCAATAATGCTAAATAGGCAAAATAGAGGTAAAGAAATAGCCAAATCATC 471
 DB 1637 AGAGAGCTCAATAATGCTAAATAGGCAAAATAGAGGTAAAGAAATAGCCAAATCATC 1696
 QY 472 TATTGCTGAGACAGACGCGGAGGACAGGATCGGGATATAACCCAGGCATTTCGAGC 531
 DB 1697 TGTGCTGACAGACAGCAGGAGGAGCAATGATCGGGATATAACCCAGGCATTTCGAGC 1756
 QY 532 CGGCAACGCGCAACCTCTTGGGTCCCTCTTGTATGGGCGCTCTGTTTTCACCTCTA 591
 DB 1757 CAGTACAGTACCTCTTGGGTCCCTCTTGTATGGGAGCTCTGT----- 1806
 QY 592 TTTCACTCTATTAAATCTTGCAACTGCAAAATAAATAAGAAA 635

Fri Feb 25 16:26:30 2005

Db	1807	CTTCACTCTATTAAATCTTGAACCTGCAAAATAAATAAGAAA	1850
Db	1457	TGTTGGCCAACTCCCAACAGCAGTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC	1516
Qy	301	AGGACTAGCTGGATTTCTTAGCCACGAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360
Db	1517	AGGAATAAAGTAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGGTGACC	1576
Qy	361	GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATC-----	411
Db	1577	GCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAA	1636
Qy	412	AGAGAGCTCACTAAATGCTTAATTAGGCAAAATAGAGGATAAGAAATAGCCAAATCATC	471
Db	1637	AGAGAGCTCACTAAATGCTTAATTAGGCAAAATAGAGGATAAGAAATAGCCAAATCATC	1696
Qy	472	TATTGCTGAGACACAGCGGGAGGACAGGATCGGATATATAACCCAGGCATTCGAGC	531
Db	1697	TGTTGCTGACGACAGCAGGAGGACAAATGATCGGATATATAACCCAGGCATTCGAGC	1756
Qy	532	CGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGATGGGCGCTCTGTTTTCACCTCA	591
Db	1757	CAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGT-----	1806
Qy	592	TTTCACTCTATTAAATCTTGCAACTGAAATAAATAAATAAATAAATAAATAAATAAATAA	635
Db	1807	CTTCACTCTATTAAATCTTGCAACTGCAAAATAAATAAATAAATAAATAAATAAATAA	1850
RESULT 35			
ABA45822			
ID	ABA45822	standard; DNA; 1894 BP.	
XX	AC	ABA45822;	
XX	DT	01-FEB-2002 (first entry)	
XX	DE	Human breast cell single exon nucleic acid probe #4517.	
XX	KW	Human; microarray; single exon probe; gene expression; breast; disease;	
XX	KW	cancer; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200157271-A2.	
XX	PD	09-AUG-2001.	
XX	PF	30-JAN-2001; 2001WO-US000662.	
XX	PR	04-FEB-2000; 2000US-0180312P.	
XX	PR	26-MAY-2000; 2000US-0207456P.	
XX	PR	30-JUN-2000; 2000US-00608408.	
XX	PR	03-AUG-2000; 2000US-00632366.	
XX	PR	21-SEP-2000; 2000US-0234687P.	
XX	PR	27-SEP-2000; 2000US-0236359P.	
XX	PR	04-OCT-2000; 2000GB-00024263.	
XX	PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX	PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX	PI	WPI; 2001-496933/54.	
XX	DR		
XX	PT	New spatially-addressable set of single exon nucleic acid probes, useful	
XX	PT	for measuring gene expression in sample derived from human breast,	
XX	PT	comprises number of single exon nucleic acid probes.	
XX	PS	Claim 1; SEQ ID NO 4517; 327pp + Sequence Listing; English.	
XX	CC	The invention relates to a spatially-addressable set of single exon	
XX	CC	nucleic acid probes for measuring gene expression in a sample derived	
XX	CC	from human breast and BT 474 cells. The method involves contacting the	
XX	CC	probes with a collection of detectably labelled nucleic acids derived	

CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match 78.7%; Score 499.6; DB 4; Length 1894;
Best Local Similarity 88.7%; Pred. No. 2.3e-144;
Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;

QY 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
DB 1217 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 1276
QY 61 CAAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
DB 1277 CAAATGCTTCTTCAATGGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGACCC 1336
QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 180
DB 1337 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 1396
QY 181 GAAATCTCAATCGACACCTCTACTACGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
DB 1397 GAAATCTCAATCGACACCTCTACTACGCCCAATTCAGCGGAGCAGTTAGAGTGGT 1456
QY 241 CATCAGCACCCTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 1457 TGTGGCCAACTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 1516
QY 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 1517 AGGATTAACCTAGATTTCTTAGGCCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACC 1576
QY 361 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCCGACCAATC----- 411
DB 1577 GCTTCCACCTTTAAACACCGGGCTTGCACTTAGCTCACACCCGACCAATCAGATACTAA 1636
QY 412 AGAGAGCTCACTAAATGCTTAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATC 471
DB 1637 AGAGAGCTCACTAAATGCTTAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATC 1696
QY 472 TATTGCTGAGAGCAGCAGCGGGAGGCAAGGATCGGGATATAAACCAGGCAATTCGAGC 531
DB 1697 TGTTCCTGACAGCAGCAGGAGGAGCAGATGATCGGGATATAAACCAGGCAATTCGAGC 1756
QY 532 CGGCAACGGCAACCCCTTTGGGTCCCTCTCTTTGATGGGCGCTCTGTTTCACTCTA 591
DB 1757 CAGCTACAGCTACCCCTCTTTGGGTCCCTCTCTTTGATGGGCGCTCTGTTTCACTCTA 1806
QY 592 TTTTCACTCTATTAATCTTCACTGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 635
DB 1807 CTTTCACTCTATTAATCTTCACTGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1850

RESULT 36
ABA25978 standard; DNA; 1894 BP.
XX ABA25978;
XX ABA25978;
XX 23-JAN-2002 (first entry)
DT

XX DE Probe #4444 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; hs.
OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
PS Claim 1; SEQ ID NO 4444; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 78.7%; Score 499.6; DB 4; Length 1894;
Best Local Similarity 88.7%; Pred. No. 2.3e-144;
Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
QY 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
DB 1217 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 1276
QY 61 CAAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
DB 1277 CAAATGCTTCTTCAATGGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGGACCC 1336
QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 180
DB 1337 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 1396
QY 181 GAAATCTCAATCGACACCTCTACTACGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
DB 1397 GAAATCTCAATCGACACCTCTACTACGCCCAATTCAGCGGAGCAGTTAGAGTGGT 1456
QY 241 CATCAGCACCCTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 1457 TGTGGCCAACTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 1516
QY 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Query Match		78.7%;	Score 499.6;	DB 4;	Length 1894;
Best Local Similarity		88.7%;	Pred. No. 2.3e-144;		
Matches 571;		Conservative	0;	Mismatches 54;	Indels 19; Gaps 2;
Qy	1	CCCTGTATCTTTAACTCTCTTGTAGTTTGTCTTCCAGAAATCAAACTGTAATACTA	60		
Db	1217	CCCTGTATCTTTAACTCTCTTGTAGTTTGTCTTCCAGAAATCGAAGCAGTAATACTA	1276		
Qy	61	CAAATTTGTTCTTCAAATGGAGCACAGATGGAGTCCATGATCAAGATCCACCGTGGACCC	120		
Db	1277	CAAATCGTCTTCAAATGGAGCCCCAGATGCGAGTCCATGATGATTAATCTTACCACGGACCC	1336		
Qy	121	CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCACCCCTCCCGAG	180		
Db	1337	CTGGACCGGCTGCTAGCCCATGCTCTGATGTTTAATGACATCAAAAGGCACCCCTCCCGAG	1396		
Qy	181	GAATCTCAATGTCACAAACCCCTTACTATGCCCAATTCAGCGGGAGCAGTTAGAGCGGT	240		
Db	1397	GAATCTCAATGTCACAAACCTCTACTACGCCCCCAATTCAGCAGAGAGCAGTTAGAGTGGT	1456		
Qy	241	CATCAGCAACCTCCCAACAGCACTTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	300		
Db	1457	TGTTGGCAACCTCCCAACAGCAGTTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	1516		
Qy	301	AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGTCACT	360		
Db	1517	AGGAATAACTAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGTCACT	1576		
Qy	361	GCATCCACCTTAACATGCGGCTTGCAACTTAGCTCAGCCCGCAATC-----	411		
Db	1577	GCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTCAGCCCGCAATCAGATACTAA	1636		
Qy	412	AGAGAGCTCACTAAATGCTAAATTAGGCAGAAATAGGAGGTAAAGAAATAGCCAAATCATC	471		
Db	1637	AGAGAGCTCACTAAATGCTAAATTAGGCAGAAATAGGAGGTAAAGAAATAGCCAAATCATC	1696		
Qy	472	TATTCCTGTAGAGCACAGCGGGAGGACAGAGATCGGGATATAAACCAGGCATTCGAGC	531		
Db	1697	TGTTGCCTGACAGCACAGCAGGAGGACAAATGATCGGGATATAAACCAGGCATTCGAGC	1756		
Qy	532	CGGCAACGGCAACCCCTTTGGGTCCTCTTGTATGGGCGCTCTGTTTTCACCTCTA	591		
Db	1757	CAGCTACAGTACCTCTTTGGGTCCTCTTGTATGGGAGCTCTGT-----	1806		
Qy	592	TTTCACCTCTATTAATCTTGCAACTGAAAAAATAAAAAA 635			
Db	1807	CTTCACCTCTATTAATCTTGCAACTGAAAAAATAAAAAATAGAAA 1850			
RESULT 38					
AAK04516					
ID	AAK04516 standard; DNA; 1894 BP.				
XX	AC AAK04516;				
XX	XX				
DT	05-NOV-2001 (first entry)				
XX	Human brain expressed single exon probe SEQ ID NO: 4507.				
DE	XX				
XX	Human; brain expressed exon; gene expression analysis; probe; microarray;				
KW	Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;				
KW	ss.				
XX	Homo sapiens.				
OS	XX				
XX	WO200157275-A2.				
PN	XX				
XX	09-AUG-2001.				
PD	XX				
XX	30-JAN-2001; 2001WO-US000667.				
PF	XX				
XX	04-FEB-2000; 2000US-0180312P.				
PR	26-MAY-2000; 2000US-0207456P.				
PR					

1517	AGGAATACTAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGTGAAC	1576
361	GCATCCACCTCTAAACATGGGGCTTGGCAACTTAGCTCACACCCGACCAATC-----	411
1577	GCTTCCACCTTTAAACACCGGGCTTGGCAACTTAGCTCACGCCCAACCAATCAGATACTAA	1636
412	AGAGAGCTCACTAAATGCTAAATTAGGCAGAAATAGGAGGTAAAGAAATAGCCAAATCATC	471
1637	AGAGAGCTCACTAAATGCTAAATTAGGCAGAAATAGGAGGTAAAGAAATAGCCAAATCATC	1696
472	TATTCCTGTAGAGCACAGCGGGAGGACAGAGATCGGGATATAAACCAGCATTCGAGC	531
1697	TGTTGCCTGACAGCACAGCAGGAGGACATGATCGGGATATAAACCAGCATTCGAGC	1756
532	CGGCAACGGCAACCCCTTTGGGTCCTTGGGTCCTTGTATGGGCGCTCTGTTTCACTCTA	591
1757	CAGCTACAGTACCTCTTTGGGTCCTTGGGTCCTTGTATGGGAGCTCTGT-----	1806
592	TTTCACTCTATTAATCTTGCAACTGAAAAAATAAAAAA 635	
1807	CTTCACTCTATTAATCTTGCAACTGAAAAAATAAAAAATAGAAA 1850	
RESULT 37		
AAK30018		
AAK30018 standard; DNA; 1894 BP.		
AAK30018;		
06-NOV-2001 (first entry)		
Human bone marrow expressed single exon probe SEQ ID NO: 4575.		
Human; bone marrow expressed exon; gene expression analysis; probe;		
microarray; cancer; leukaemia; lymphoma; myeloma; ss.		
Homo sapiens.		
WO200157276-A2.		
09-AUG-2001.		
30-JAN-2001; 2001WO-US000668.		
04-FEB-2000; 2000US-0180312P.		
26-MAY-2000; 2000US-0207456P.		
30-JUN-2000; 2000US-00608408.		
03-AUG-2000; 2000US-00632366.		
21-SEP-2000; 2000US-0234687P.		
27-SEP-2000; 2000US-0236359P.		
04-OCT-2000; 2000GB-00024263.		
(MOLE-) MOLECULAR DYNAMICS INC.		
Penn SG, Hanzel DK, Chen W, Rank DR;		
WPI; 2001-488900/53.		
Human genome-derived single exon nucleic acid probes useful for analyzing		
gene expression in human bone marrow.		
Example 4; SEQ ID NO 4575; 658pp + Sequence Listing; English.		
The present invention provides a number of single exon nucleic acid		
probes which are derived from genomic sequences expressed in the human		
bone marrow. They can be used to measure gene expression in bone marrow		
samples, which may enable the improved diagnosis and treatment of cancers		
such as lymphoma, leukaemia and myeloma. The present sequence is one of		
the probes of the invention		
Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;		

PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 XX brains.
 XX
 PS Example 4; SEQ ID NO 4507; 650pp + Sequence Listing; English.
 XX
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention
 XX
 XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 XX
 SQ
 Query Match 78.7%; Score 499.6; DB 4; Length 1894;
 Best Local Similarity 88.7%; Pred. No. 2.3e-144;
 Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTTGTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
 DB 1217 CCTGTATCTTTAACTCTTGTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 1276
 QY 61 CAAATCTCTTCAAAATGGAGCAGATGAGTCCATGATCAAGATCCACCGTGGACCC 120
 DB 1277 CAAATCTCTTCAAAATGGAGCAGATGAGTCCATGATCAAGATCCACCGTGGACCC 1336
 QY 121 CTGACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTTAGAGCAGCCCTCCGAG 180
 DB 1337 CTGACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTTAGAGCAGCCCTCCGAG 1396
 QY 181 GAAATCTCACTGACACACCTTACTATGCCCCCAATTCAGCGGAGCAGTATAGAGGCT 240
 DB 1397 GAAATCTCACTGACACACCTTACTATGCCCCCAATTCAGCGGAGCAGTATAGAGGCT 1456
 QY 241 CATCAGGCAACTCTCCCAACAGCACTTGGGTTTCTCTGTTGAGAGGGGAGCTGAGAGAC 300
 DB 1457 TGTGGCCAACTCTCCCAACAGCACTTGGGTTTCTCTGTTGAGAGGGGAGCTGAGAGAC 1516
 QY 301 AGGACTAGCTGATTTCTTAGGCCAACAGAAATCCCTAGCCCTAGCTGGGAGGTGACT 360
 DB 1517 AGGAAATTAAGTATTTCTAGACCAACTAAGAAATCCCTAGCTAGCTGGGAGGTGACT 1576
 QY 361 GCATCCACCTTAACATGCGGCTTGCACCTAGCTACACCCGACCAATC----- 411
 DB 1577 GCTTCCACCTTTAAACACCGGGCTTGCACCTAGCTACACCCGACCAATCAGATATAA 1636
 QY 412 AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATC 471
 DB 1637 AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATC 1696
 QY 472 TATTGCTGAGACACAGCGGAGGGAAGAGATCGGGATATAAACCCAGGCAATTCGAGC 531
 DB 1697 TGTTCCTGACAGACAGCAGGAGGGAAGATGATCGGGATATAAACCCAGGCAATTCGAGC 1756
 QY 532 CGGCAACGGCAACCCCTTTGGGTCCCTTCTTGTATGGCGCTCTGTTTCACTCTA 591
 DB 1757 CAGCTACAGCTACCTCTTTGGGTCCCTTCTTGTATGGCGCTCTGTTTCACTCTA 1806
 QY 592 TTTCACTCTATTAAATCTTGCAACTGAAATAAAAAAAAAAAAAA 635

DB 1807 CTTCACTCTTAAATCTTTGCAACTGCAAAATAAATAATAGAAA 1850
 RESULT 39
 ABS29670
 ID ABS29670 standard; DNA; 1894 BP.
 XX
 AC ABS29670;
 XX
 XX 25-FEB-2003 (first entry)
 XX
 DE Human liver single exon probe, SEQ ID No 4660.
 XX
 XX Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157273-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US000664.
 PF
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX WPI; 2001-488898/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 PT
 XX Claim 1; SEQ ID NO 4660; 658pp; English.
 XX
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABS25011-ABS1005 represent human
 CC liver single exon nucleic acid probes of the invention. Note: The
 CC sequence information for this patent does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 Query Match 78.7%; Score 499.6; DB 4; Length 1894;
 Best Local Similarity 88.7%; Pred. No. 2.3e-144;
 Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTTGTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
 DB 1217 CCTGTATCTTTAACTCTTGTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 1276
 QY 61 CAAATCTCTTCAAAATGGAGCAGATGAGTCCATGATCAAGATCCACCGTGGACCC 120
 DB 1277 CAAATCTCTTCAAAATGGAGCAGATGAGTCCATGATCAAGATCCACCGTGGACCC 1336

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

QY 121 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 180
Db 1337 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 1396
QY 181 GAAATCTCAATGCAACAACCTTACTATGCCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
Db 1397 GAAATCTCAATGCAACAACCTTACTATGCCCCAATTCAGCGGAGCAGTTAGAGTGGT 1456
QY 241 CATGACCAACTCCCAACAGCACTTGGGTTTCTGTTGAGGGGGGAGTCTGAGAGAC 300
Db 1457 TGTGGCCCACTCCCAACAGCACTTGGGTTTCTGTTGAGGGGGGAGTCTGAGAGAC 1516
QY 301 AGGACTAGCTGGATTTCTTAGGCCAAAGAAATCCCTAAGCTTACTGCTGGGAGTGAAT 360
Db 1517 AGGATAAATAGCTTTCTTAGGCCAAAGAAATCCCTAAGCTTACTGCTGGGAGTGAAT 1576
QY 361 GATCCACCTTAACATGGGCTTGGCACTTAGCTTACACCCGACCAATC----- 411
Db 1577 GCTTCCACCTTTAAACACCGGGCTTGGCACTTAGCTTACACCCGACCAATCAGATACTAA 1636
QY 412 AGAGAGCTCACTAAATGCTAATTAGGCCAAAGAAATAGGAGGTAAAGAAATAGCCAAATCATC 471
Db 1637 AGAGAGCTCACTAAATGCTAATTAGGCCAAAGAAATAGGAGGTAAAGAAATAGCCAAATCATC 1696
QY 472 TATTGCTGAGACACAGCGGAGGGAACAAGGATCGGGATATAAACCCAGGCAATTCGAGC 531
Db 1697 TGTGGCTGACACACAGCGGAGGGAACAAGGATCGGGATATAAACCCAGGCAATTCGAGC 1756
QY 532 CGGCAACCGCAACCCCTTTGGTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTA 591
Db 1757 CAGCTACAGCTACCTCTTTGGTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTA 1806
QY 592 TTTCACTCTATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 635
Db 1807 CTTCACTCTATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 1850

RESULT 40

AAI04422

ID AAI04422 standard; DNA; 1894 BP.

XX AAI04422;

XX 09-OCT-2001 (first entry)

XX Probe #4413 used to measure gene expression in human breast sample.
XX Probe; human; breast disease; breast cancer; development disorder; ss;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX WO200157270-A2.

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US0000661.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression in
PT

PT a human breast.

XX Claim 25; SEQ ID NO 4413; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 78.7%; Score 499.6; DB 5; Length 1894;

Best Local Similarity 88.7%; Pred. No. 2.3e-144;

Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;

QY 1 CCCTGTATCTTTAACTCTCTTGTAGTTTCTTCCAGAAATCGAAGCAGTAAACTA 60
Db 1217 CCCTGTATCTTTAACTCTCTTGTAGTTTCTTCCAGAAATCGAAGCAGTAAACTA 1276
QY 61 CAATTTCTTCTCAAAATGGAGCACCAGATGGAGTCCATGATTAAGATCCACCGTGGACCC 120
Db 1277 CAATTCGTCTTCAAAATGGAGCACCAGATGGAGTCCATGATTAAGATCCACCGTGGACCC 1336
QY 121 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 180
Db 1337 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 1396
QY 181 GAAATCTCAATGCAACAACCTTACTATGCCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
Db 1397 GAAATCTCAATGCAACAACCTTACTATGCCCCAATTCAGCGGAGCAGTTAGAGTGGT 1456
QY 241 CATGACCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGAGTCTGAGAGAC 300
Db 1457 TGTGGCCCACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGAGTCTGAGAGAC 1516
QY 301 AGGACTAGCTGGATTTCTTAGGCCAAAGAAATCCCTAAGCTTACTGCTGGGAGTGAAT 360
Db 1517 AGGATAAATAGCTTTCTTAGGCCAAAGAAATCCCTAAGCTTACTGCTGGGAGTGAAT 1576
QY 361 GCATCCACCTTAACATGGGCTTGGCACTTAGCTTACACCCGACCAATC----- 411
Db 1577 GCTTCCACCTTTAAACACCGGGCTTGGCACTTAGCTTACACCCGACCAATCAGATACTAA 1636
QY 412 AGAGAGCTCACTAAATGCTAATTAGGCCAAAGAAATAGGAGGTAAAGAAATAGCCAAATCATC 471
Db 1637 AGAGAGCTCACTAAATGCTAATTAGGCCAAAGAAATAGGAGGTAAAGAAATAGCCAAATCATC 1696
QY 472 TATTGCTGAGACACAGCGGAGGGAACAAGGATCGGGATATAAACCCAGGCAATTCGAGC 531
Db 1697 TGTGGCTGACACACAGCGGAGGGAACAAGGATCGGGATATAAACCCAGGCAATTCGAGC 1756
QY 532 CGGCAACCGCAACCCCTTTGGTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTA 591
Db 1757 CAGCTACAGCTACCTCTTTGGTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTA 1806
QY 592 TTTCACTCTATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 635
Db 1807 CTTCACTCTATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 1850

Search completed: February 21, 2005, 07:45:38

Job time : 339.546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:54:34 ; Search time 111.885 Seconds
(without alignments)
9286.612 Million cell updates/sec

Title: US-09-319-156B-6
Perfect score: 635
Sequence: 1 cccgtatctttaaactctct.....tgaaaaa.....635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	635	100.0	635	4	US-08-979-847B-102
2	545	85.8	1329	4	US-08-979-847B-108
3	533.8	84.1	2946	3	US-09-175-928-3
4	516.2	81.3	13537	4	US-09-949-016-17382
5	515.8	81.2	2763	4	US-09-949-016-5640
6	510.6	80.4	145320	4	US-09-949-016-15858
7	453	71.3	99580	4	US-09-949-016-17411
8	448	70.6	168394	4	US-09-949-016-13002
9	442.6	69.7	77772	4	US-09-949-016-17417
10	442.6	69.7	77997	4	US-09-949-016-12249
11	405.4	63.8	601	4	US-09-949-016-44300
12	367.2	57.8	149971	4	US-09-949-016-13590
13	292	46.0	245286	4	US-09-949-016-15497
14	291.6	45.9	601	4	US-09-949-016-133739
15	288	45.4	194790	4	US-09-949-016-15393
16	273.2	43.0	601	4	US-09-949-016-44296
17	269	42.4	219964	4	US-09-949-016-15086
18	262.8	41.4	780	4	US-09-573-080A-385
19	252.2	39.7	601	4	US-09-949-016-103638
20	252.2	39.7	601	4	US-09-949-016-103639
21	251	39.5	601	4	US-09-949-016-103637
22	251	39.5	256287	4	US-09-949-016-14608
23	249.4	39.3	601	4	US-09-949-016-44299
24	245.8	38.7	84571	4	US-09-949-016-17420
25	245.6	38.7	140224	4	US-09-949-016-17002
26	239.4	37.7	89584	4	US-09-949-016-17068
27	233.6	36.8	279	1	US-08-686-878A-50

28	233.6	36.8	279	1	US-08-721-489-4	Sequence 4, Appli
29	219.6	34.6	8523	4	US-09-573-080A-21	Sequence 21, Appli
30	196.6	31.0	456	4	US-09-621-976-9366	Sequence 9366, Ap
31	195.6	30.8	601	4	US-09-949-016-128750	Sequence 128750,
32	180.4	28.4	154023	4	US-09-949-016-17057	Sequence 17057, A
33	178	28.0	601	4	US-09-949-016-44301	Sequence 44301, A
34	176.8	27.8	601	4	US-09-949-016-195792	Sequence 195792, A
35	176.8	27.8	251672	4	US-09-949-016-17296	Sequence 17296, A
36	176.8	27.8	251682	4	US-09-949-016-11973	Sequence 11973, A
37	175	27.6	601	4	US-09-949-016-44310	Sequence 44310, A
38	174.6	27.5	116652	4	US-09-949-016-13413	Sequence 13413, A
39	174	27.4	39686	4	US-09-949-016-13633	Sequence 13633, A
40	174	27.4	49487	4	US-09-949-016-11770	Sequence 11770, A
41	166.8	26.3	131631	4	US-09-949-016-11757	Sequence 11757, A
42	164.8	26.0	57507	4	US-09-949-016-15019	Sequence 15019, A
43	156	24.6	64291	4	US-09-949-016-16278	Sequence 16278, A
44	156	24.6	117410	4	US-09-949-016-12262	Sequence 12262, A
45	147.2	23.2	601	4	US-09-949-016-202370	Sequence 202370,

ALIGNMENTS

RESULT 1

US-08-979-847B-102
; Sequence 102, Application US/08979847B
; Patent No. 6582703

GENERAL INFORMATION:

APPLICANT: PERRON, HERVE
BESEME, FREDERIC
BEDIN, FREDERIC
PARANHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TURE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE, PLC

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER:

APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-No. 6582703-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BERRIDGE, WILLIAM P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 39046A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 635 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-08-979-847B-102

Query Match 100.0%; Score 635; DB 4; Length 635;
Best Local Similarity 100.0%; Pred. No. 2e-203;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAATACTA 60
DB 1 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAATACTA 60

QY 61 CAAATGTTCTTCAATGAGACACAGATGGAGTGCATGACTTAAGATCCACCGTGGACCC 120
DB 61 CAAATGTTCTTCAATGAGACACAGATGGAGTGCATGACTTAAGATCCACCGTGGACCC 120

QY 121 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 121 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

QY 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300

QY 301 AGGACTAGCTGGATTCCTAGGCCACGAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 301 AGGACTAGCTGGATTCCTAGGCCACGAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCGGCAATTCAGAGGCTC 420
DB 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCGGCAATTCAGAGGCTC 420

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCCCAATCATATTGCCTG 480
DB 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCCCAATCATATTGCCTG 480

QY 481 AGAGCACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCAGCGCGCAACGG 540
DB 481 AGAGCACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCAGCGCGCAACGG 540

QY 541 CAACCCCTCTGGGTCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 541 CAACCCCTCTGGGTCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 600

QY 601 ATTAATCTTGCACCTGAAAAA 635
DB 601 ATTAATCTTGCACCTGAAAAA 635

RESULT 2
US-08-979-847B-108
; Sequence 108, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUXE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA

COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-No. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-08-979-847B-108

Query Match 85.8%; Score 545; DB 4; Length 1329;
Best Local Similarity 92.4%; Pred. No. 7.2e-173;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAATACTA 60
DB 257 CCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTTCCAGAAATGAAGCTGTAAGCTA 316

QY 61 CAAATGTTCTTCAAAATGGAGCACAGATGGAGTGCATGACTTAAGATCCACCGTGGACCC 120
DB 317 CAAATGTTCTTCAAAATGGAAACCCAGATGCAGTCCATGACTTAAATCTACCGTGGACCC 376

QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 377 CTGACCGGCTCTAGACTATGCTCTGATGTTAATGACATTAAGTCAACCCCTCCCGAG 436

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 437 GAAATCTCAACTGCACAAACCCCTACTACACTCCCAATTCAGTAGGAGCAGTTAGAGCAGT 496

QY 241 CATCAGCCAACTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 497 TGTGAGCCAACTCCCAACAGACTTGGGTTTCTGTTGAGAGGGTGGACTGAGAGAC 556

QY 301 AGGACTAGCTGGATTCCTAGGCCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 557 AGGACTAGCTGGATTCCTAGGCTGACTAAGAAATCCCAAGCCTAATCTGGGAAGGTGACC 616

QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGCAATTCAGAGAGCTC 420
DB 617 GCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGCAATTCAGAGAGCTC 676

QY 421 ACTAAATGCTAATTAGGCAAAATAGAGGTAAGAAATAGCCCAATCATATTGCCTG 480
DB 677 ACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCCAATCATATTGCCTG 736

QY 481 AGAGCACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCAGCGCGCAACGG 540
DB 737 AGAGCACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCAGCGCGCAACAG 796

QY 541 CAAACCCCTTTGGGTCCTCCCTCTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 797 CAAACCCCTTTGGGTCCTCCCTCTGATGGAGCTCTGTTTCACTCTATTTCACCTCT 856

QY 601 ATTAATCTTGCACCTGAA 619

Db 857 ATTAATCATGCAACTGCA 875
|||||

RESULT 3

US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B A17172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3

Query Match 84.1%; Score 533.8; DB 3; Length 2946;
Best Local Similarity 91.5%; Pred. No. 7.2e-169;
Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;
QY 1 CCTGTATCTTTAACTCTTGTAAAGTTGCTCTTCAGAAATCAAACTGTAAACTA 60
Db 2323 CCTGTATCTTTAACTCTTGTAAAGTTGCTCTTCAGAAATCAAACTGTAAACTA 2382
QY 61 CAAATTGTTCTTCAATGGAGCCAGATGAGTCCATGACTAAGATCCACCGGTGACCC 120
Db 2383 -----CAAAATGGAGCCAGATGAGTCCAAAGATTAAGATCTACCGCAGACCC 2430
QY 121 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG 180
Db 2431 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG 2490
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 2491 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 2550
QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 2551 CGTCGGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 2610
QY 301 AGGACTAGCTGGATTTCCTAGGCAACAGAGATCCCTAAGACCTAGCTGGGAAGGTGACT 360
Db 2611 AGGACTAGCTGGATTTCCTAGGCTGACTAAGATCCCTAAGACCTAGCTGGGAAGGTGACT 2670
QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACAACCCGACCAATCAGAGAGCTC 420
Db 2671 ACATCCACCTTTAAACACAGGGCTTGCAACTTAGCTCACAACCTGACCAATCAGAGAGCTC 2730
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCTG 480
Db 2731 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCTG 2790
QY 481 AGAGCAGCGGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCAGCGCGCAACCG 540
Db 2791 AGAGCAGCGGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCAGCGCGCAACCG 2850
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCT 600
Db 2851 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCT 2910

QY 601 ATTAATCTTGCACACTGAAAAAAGAAAAA 635
Db 2911 ATTAATCTTGCACACTGCAAAAAAAGAAAAA 2945
|||||

RESULT 4

US-09-949-016-17382
; Sequence 17382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17382
; LENGTH: 13537
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17382

Query Match 81.3%; Score 516.2; DB 4; Length 13537;
Best Local Similarity 91.1%; Pred. No. 1.7e-162;
Matches 564; Conservative 0; Mismatches 43; Indels 12; Gaps 1;
QY 1 CCTGTATCTTTAACTCTTGTAAAGTTGCTCTTCAGAAATCAAACTGTAAACTA 60
Db 10931 CCTGTATCTTTAACTCTTGTAAAGTTGCTCTTCAGAAATCAAACTGTAAACTA 10990
QY 61 CAAATTGTTCTTCAATGGAGCCAGATGAGTCCATGACTAAGATCCACCGGTGACCC 120
Db 10991 -----CAAAATGGAGCCAGATGAGTCCAAAGATTAAGATCTACCGCAGACCC 11038
QY 121 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG 180
Db 11039 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGAGCCACCCCTCCCGAG 11098
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 11099 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 11158
QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 11159 CGTCGGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 11218
QY 301 AGGACTAGCTGGATTTCCTAGGCAACAGAGATCCCTAAGACCTAGCTGGGAAGGTGACT 360
Db 11219 AGGACTAGCTGGATTTCCTAGGCTGACTAAGATCCCTAAGACCTAGCTGGGAAGGTGACT 11278
QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACAACCCGACCAATCAGAGAGCTC 420
Db 11279 ACATCCACCTTTAAACACAGGGCTTGCAACTTAGCTCACAACCTGACCAATCAGAGAGCTC 11338
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCTG 480
Db 11339 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCTG 11398
QY 481 AGAGCAGCGGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCAGCGCGCAACCG 540
Db 11399 AGAGCAGCGGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCAGCGCGCAACCG 11458
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCT 600
|||||

Db 11459 CAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTCATGCTATTTCACCTCT 11518

Qy 601 ATTAATCTTGCAACTGAA 619

Db 11519 ATTAATCTTGCAACTGCA 11537

RESULT 5

US-09-949-016-5640

; Sequence 5640, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5640

; LENGTH: 2763

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-5640

Query Match 81.2%; Score 515.8; DB 4; Length 2763;

Best Local Similarity 91.2%; Pred. No. 8.1e-163;

Matches 563; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

Qy 1 CCTGTATCTTTAACTCTCTTGTAGTTTGTCTCTCCAGAACTCAAACTGTAAACTA 60

Db 2158 CCTGTATCTTTAACTCTCTTGTAACTTTGTCTCTCCAGAACTCAAACTGTAAACTA 2217

Qy 61 CAAATTTGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Db 2218 -----CAATGGAGCCCAAGATGCAGTCCAGCTAAGATCTACCGGAGACC 2265

Qy 121 CTGGACGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 180

Db 2266 CTGGACGGCTCTAGCCCATGCTGATGTTAATGACATCAAGGCAACCCCTCCCGAG 2325

Qy 181 GAAATCTCACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240

Db 2326 GAAATCTCACTGCACAAACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGACGGT 2385

Qy 241 CATCAGCCAACTCTCCCAACAGACCTTGGGTTTTCTGTGAGAGGGGGGACTGAGAGAC 300

Db 2386 CGTCGGCAACCTCTCCCAACAGACCTTAGGTTTTCTGTGAGATGGGGGACTGAGAGAC 2445

Qy 301 AGACTAGCTGGATTTCTTAGGCAACAGAGATCCCTAAGCCTAGCTGGAGGTGACT 360

Db 2446 AGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGAGGTGACC 2505

Qy 361 GCATCCACCTCTAAACATGGGGTTGCAACTTAGCTCTACACCCGACCAATTCAGAGAGCTC 420

Db 2506 ACATCCACCTTTAAACAGGGGCTTGCACTTAGCTCTACACCTGACCAATTCAGAGAGCTC 2565

Qy 421 ACTAAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTTG 480

Db 2566 ACTAAAATGCTAATTAGGCAAAAAGAGGTTAAAGAAATAGCAATCATCTATTGCTTG 2625

Qy 481 AGAGCAGCGGGAGGAGGACAGGATCGGGATATAACCCAGGATTCGAGCCGGCAACGG 540

Db 2626 AGAGCAGCGGGAGGAGGACAGGATCGGGATATAACCCAGGATTCGAGCCGGCAACGG 2685

Qy 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTTCACCTATTTCACCTCT 600

Db 2686 CAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT 2745

Qy 601 ATTAATCTTGCAACTG 617

Db 2746 ATTAATCTTGCAACTG 2762

RESULT 6

US-09-949-016-15858/c

; Sequence 15858, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15858

; LENGTH: 145320

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-15858

Query Match 80.4%; Score 510.6; DB 4; Length 145320;

Best Local Similarity 90.0%; Pred. No. 6e-160;

Matches 575; Conservative 0; Mismatches 49; Indels 15; Gaps 2;

Qy 1 CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAACTCAAACTGTAAAGCTA 60

Db 92601 CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAACTCAAAAGTTGTAAGCTA 92542

Qy 61 CAAATTTGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120

Db 92541 CAAATCGTTCTTCAAAATGGAAACCCAGATGAAGTCCATGACTAAGATCTACCGTGGACCC 92482

Qy 121 CTGGACGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 180

Db 92481 CTGGACGGCTCTAGCCCATGCTCCCAATTTGTAATGATATCGAACGCAACCCCTCCCGAG 92422

Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240

Db 92421 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 92362

Qy 241 CATCAGCCAACTCTCCCAACAGACCTTGGGTTTTCTGTGAGAGGGGGGACTGAGAGAC 300

Db 92361 CGTCAGCCAACTCTCCCAACAGACCTTGGGTTTTCTGTGAGTGGGGGACTGAGAGAC 92302

Qy 301 AGACTAGCTGGATTTCTTAGGCAACAGAGATCCCTAAGCCTAGCTGGGAAAGGTGACT 360

Db 92301 AGACTAGCTGGATTTCTTAGGCAACAGAGATCCCTAAGCCTAGCTGGGAAAGGTGACC 92242

Qy 361 GCATCCACCTCTAAACATGGGGTTGCAACTTAGCTCTACACCCGACCAATC----- 411

Db 92241 ACATCCACCTTTAAACATGGGGCTTGCACCTTAGCTCTACACCCGACCAATCAGGTAGTAA 92182

Qy 412 AGAGAGCTCACTAAAATGCTAATTTAGGCAAAAATAGGAGGTAAAGAAATAGCAATCATCTC 471

Db 92181 AGAGAGCTCACTAAAATGCTAATTTAGCAAAAAGAGGAGTTAAAATAATAGCAATCATCTC 92122

Qy 472 TATTGCCTGAGAGCAGCGGGAGGACAAAGGATCGGGATATAACCCAGGATTCGAGCCGGCAACGG 531

Db 92121 TATTGCCTGAGAGCAGCGGGAGGACAAAGGATCGGGATATAACCCAGGATTCGAGCCGGCAACGG 92062

QY 532 CGGCAACGGCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACTCTA 591
 Db 92061 CGGCAACGGCTACCTTTTGGGTCCCTCTTGTATGGGGCTCTCTCT-----GT 92008
 QY 592 TTTCACTCTATTAAATCTTCAACTGAAAAAATA 630
 Db 92007 CTTCACTCTATTAAATCTTCAACTGAAAAAATA 91969

RESULT 7

US-09-949-016-17411/c
 ; Sequence 17411, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17411
 ; LENGTH: 99580
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-17411

Query Match 71.3%; Score 453; DB 4; Length 99580;
 Best Local Similarity 88.9%; Pred. No. 1.2e-140;
 Matches 538; Conservative 0; Mismatches 55; Indels 12; Gaps 4;

QY 1 CCCTGTATCTTTAACTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 60
 Db 6780 CCCTGTATCTTTAACTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 6721
 QY 61 CAAATTTGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 6720 CAAATTTGTTCTTCAAAATAGAACCCAGATGAGTCCATGACTAAGATCTACCGTGGACCC 6661
 QY 121 CTGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATGGAAGGACCCCTCCCGAG 180
 Db 6660 CTGGAACAGCCTGCTAGCCCATGCTCCGATGTTAATGACATGGAAGGACCTCTCTCTGAT 6601
 QY 181 GAAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTT-AGAGGG 239
 Db 6600 GAAATCTCAACTGACACGACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAAGAGCGAG 6541
 QY 240 TCATCAGCAACCTCCCAACAGACACTTGGGTTTTCTTGTGAGAGGGGAGCTGAGAGA 299
 Db 6540 TCGTGGCCCACTCCCAACAGACACTTGGGTTTTCTTGTGAGAGGGGAGCTGAGAGA 6481
 QY 300 CAGGACTAGCTGGATTTCTTAGGCGCAAGAAATCCCTAAGCCTAGCTGGG-AAAGTGA 358
 Db 6480 CAGGACTAGCTGGATTTCTTAGGCGCAAGAAATCCCTAAGCCTAGCTGGGAAAGTGA 6421
 QY 359 CTGATCCACCTCTTAAACATGGGGTTTGCACATTTAGCTCAGCCGACCAATC----- 411
 Db 6420 CCGCACCTACCTTTTAAACATGGGGTTTGCACATTTAGCTCAGCCGACCAATCAGGTAGT 6361
 QY 412 --AGAGGCTCACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAG-CCAATC 468
 Db 6360 AAAGGGGCTCACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGTCAATC 6301
 QY 469 ATCTATTCCTGAGAGCACAGGGGAGGACAGGATCGGATATATAAACCAGGCAATTCG 528
 Db 6300 ATATATCGCTGAGAGCACAGGGGAGGACAGGATCGGATATATAAACCAGGCAATTCG 6241

QY 529 AGCCGCAACGGCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACT 588
 Db 6240 AGCAGGAGCAGCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACT 6181
 QY 589 CTATT 593
 Db 6180 CTATT 6176

RESULT 8

US-09-949-016-13002/c
 ; Sequence 13002, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13002
 ; LENGTH: 168394
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(168394)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-949-016-13002

Query Match 70.6%; Score 448; DB 4; Length 168394;
 Best Local Similarity 88.7%; Pred. No. 8.4e-139;
 Matches 534; Conservative 0; Mismatches 55; Indels 13; Gaps 4;

QY 1 CCCTGTATCTTTAACTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 60
 Db 25264 CCCTGTATCTTTAACTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 25205
 QY 61 CAAATTTGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 25204 CAAATTTGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCTTACTCGAGACCC 25145
 QY 121 CTGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 180
 Db 25144 --GGACGAGCTGCTAGCCCATGCTCC-AGTTAATAACATCAAGGACCTCTCCAGAG 25088
 QY 181 GAAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 Db 25087 GAAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCAAT 25028
 QY 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTTGTGAGAGGGGAGCTGAGAGAC 300
 Db 25027 CATCGGCAATCT-CTCAACAGCAGCTTGGGTTTCTTGTGAGAGGGGAGCTGAGAGAC 24969
 QY 301 AGGACTAGCTGGATTTCTTAGGGCCACGAAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 24968 AGGACTAGCTGGATTTCTTAGGGCCACGAAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACC 24909
 QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 411
 Db 24908 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCTAGCTGGGAAGGTGAA 24849
 QY 412 AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATC 471

Db 24848 AGAGAGTTCACTAAATGCTAACTAGGCAAAACAGGAGGTAAAGAAATAGCCAAATCATC 24789

Qy 472 TATTCCTGAGACACACGGGAGGACAAAGATCGGATATAAACCCAGGCAATTCGAGC 531

Db 24788 TATTCCTGAGACACAGGGGAGGATATGATCAGATATATACCCAGGCAATTCGAGC 24729

Qy 532 CGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTAAGGCGCTCTGTTTCACTCTA 591

Db 24728 TGGCAATGGCTACCCCTTTGGGTCCCTCCCTTTGTAAGGCGCTCTGTTTCACTCTA 24669

Qy 592 TT 593

Db 24668 TT 24667

RESULT 9

US-09-949-016-17417/c

; Sequence 17417, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17417

; LENGTH: 77772

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(77772)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17417

Query Match 69.7%; Score 442.6; DB 4; Length 77772;

Best Local Similarity 88.9%; Pred. No. 3.4e-137;

Matches 538; Conservative 0; Mismatches 54; Indels 13; Gaps 5;

Qy 1 CCCTGTATCTTTAACTCCTTGTGTTAACTTTGTTCTCTCCAGAAATCAAACTGTAAACTA 60

Db 15972 CCCTGTATCTTTAACTCCTTGTGTTAACTTTGTTCTCTCCAGAAATCAAACTGTAAACTA 15913

Qy 61 CAAATTTGTTCTTCAAAATGGAGCACACAGATGGAGTCCATGATCAAGATCCACCGTGACCC 120

Db 15912 CAAA-TGTTCTTCAAAATGGAGCCCCAGATGCAGTCCATGATCAAGATCTACCGTGACCC 15854

Qy 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTAAATGACATTTAAGCACCCCTCCCGAG 180

Db 15853 CTGACCTGGCTGTAGCCCATGCTCCAAATGTTGATGACATCAAGGACCACTCCCGAG 15794

Qy 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240

Db 15793 GAAATCTCAACTGCATGACCTTACTAGCCCCCAATTCAGCAGGAAGCAGTTACAGCGT 15734

Qy 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGACTGAGAGAC 300

Db 15733 CATCGGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGACTTAAAGAGAC 15674

Qy 301 AGGACTAGCTGATTTCTTAGGCCCAACGAGAATCCCTTAAGCCCTAGCT-GGGGAAGTGAC 359

Db 15673 AGGACTAGCTGATTTCTTAGGCCGACCTAAGAAATTTCTTAAGCCCTAGCTGGGGAAGTGAC 15614

Qy 360 TGCATCCACCTCTAAACATATGGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 411

Db 15613 CACACCCACCTTTAAACATGGGGCTTGTAACTCAGCTCACACAGACCAATCAGGTAGTA 15554

Qy 412 -AGAGAGCTCACTAAATGCTAATTAGGC-AAATAAGGAGGTAAAGAAATAGCCA-ATC 468

Db 15553 AAGAGAGCTCACTAAATGCTAATTAGGC-AAATAAGGAGGTAAAGAAATAGCCA-ATC 15494

Qy 469 ATCTATTGCTGAGACACAGCGGAGGACAGGATCGGATATAAAACCCAGGCAATTCG 528

Db 15493 ATCTGCTGCTGAGACACAGCGGAGGAGGACATGATCAGGATATAAAACCCAGGCAATTC 15434

Qy 529 AGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGGCGCTCTGTTTCACT 588

Db 15433 AGCGGAATCGGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTCACT 15374

Qy 589 CTATT 593

Db 15373 ATATT 15369

RESULT 10

US-09-949-016-12249/c

; Sequence 12249, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12249

; LENGTH: 77997

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(77997)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-12249

Query Match 69.7%; Score 442.6; DB 4; Length 77997;

Best Local Similarity 88.9%; Pred. No. 3.4e-137;

Matches 538; Conservative 0; Mismatches 54; Indels 13; Gaps 5;

Qy 1 CCCTGTATCTTTAACTCCTTGTGTTAACTTTGTTCTCTCCAGAAATCAAACTGTAAACTA 60

Db 15976 CCCTGTATCTTTAACTCCTTGTGTTAACTTTGTTCTCTCCAGAAATCAAACTGTAAACTA 15917

Qy 61 CAAATTTGTTCTTCAAAATGGAGCACACAGATGGAGTCCATGATCAAGATCCACCGTGACCC 120

Db 15916 CAAA-TGTTCTTCAAAATGGAGCCCCAGATGCAGTCCATGATCAAGATCTACCGTGACCC 15858

Qy 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTAAATGACATTTAAGCACCCCTCCCGAG 180

Db 15857 CTGACCTGGCTGTAGCCCATGCTCCAAATGTTGATGACATCAAGGACCACTCCCGAG 15798

Qy 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240

Db 15797 GAAATCTCAACTGCATGACCTTACTAGCCCCCAATTCAGCAGGAAGCAGTTACAGCGT 15738

Qy 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGACTGAGAGAC 300

Db 15737 CATCGGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGTAATAAGAGAC 15678

QY 301 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGAAATCCCTAAGCCTAGCT-GGGAAGGTGAC 359
 DB 15677 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGAAATCCCTAAGCCTAGCTGGGAAGGTGAC 15618
 QY 360 TGCATCCACTTAAACATGGGCTTGCAACTTGAATCTAGCTCACCAGCAATC----- 411
 DB 15617 CACACCCACCCTTAAACATGGGCTTGCAACTTGAATCTAGCTCACCAGCAATCAGGTAGTA 15558
 QY 412 -AGGAGCTCCTAAATATGCTAATAGGC-AAAAATAGGAGGTAAAGAAATATGCCA-ATC 468
 DB 15557 AAGAGAGCTCCTAAATATGCTAATAGGC-AAAAATAGGAGGTAAAGAAATATGCCAATCAGATC 15498
 QY 469 ATCTATTGCTGAGAGCAGACAGCGGGAGGAGCAAGGATCGGGATATATAACCCAGGCATTTCG 528
 DB 15497 ATCTGCTGCTGAGAGCAGACAGCGGGAGGAGCAATGATCAGATATATAACCCAGGCATTTC 15438
 QY 529 AGCGGCAACGGCAACCCCTTTGGGTCCCTCTCTTTGATGGGCTCTGTTTCACT 588
 DB 15437 AGCGGAATCGGCAACCCCTTTGGGTCCCTCTCTTTGATGGGCTCTGTTTCACT 15378
 QY 589 CTATT 593
 DB 15377 ATATT 15373

RESULT 11
 US-09-949-016-44300/c
 ; Sequence 44300, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 44300
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 ; ORIGIN: 016-44300

Query Match 63.8%; Score 405.4; DB 4; Length 601;
 Best Local Similarity 88.1%; Pred. No. 5e-126;
 Matches 489; Conservative 1; Mismatches 52; Indels 13; Gaps 4;
 QY 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAACTA 60
 DB 552 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAACTA 493
 QY 61 CAAATTTGTTCTTAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 DB 492 CAAATTTGTTCTTAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 433
 QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGCACCCTCCCGAG 180
 DB 432 -GGACACCGCTCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGCACCCTCCCGAG 376
 QY 181 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 DB 375 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCAAT 316
 QY 241 CATCAGCAACCTCCCAACAGCACTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
 DB 315 CATCGGCCAATCT-CYCAACAGCACTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 257

QY 301 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 DB 256 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACC 197
 QY 361 GCATCCACTCTAAACATGGGCTTGCAACTTGAATCTAGCTCACCAGCAATC----- 411
 DB 196 GCATCCACTCTTAAACATGGGCTTGCAACTTGAATCTAGCTCACCAGCAATCAGGTAGTAA 137
 QY 412 AGAGAGCTCCTAAATATGCTAATAGGC-AAAAATAGGAGGTAAAGAAATAGCCAATCATC 471
 DB 136 AGAGAGTCTCCTAAATATGCTAATAGGC-AAAAATAGGAGGTAAAGAAATAGCCAATCATC 77
 QY 472 TATTGCTGAGAGCAGACAGCGGGAGGAGCAAGGATCGGGATATATAACCCAGGCATTTCGAGC 531
 DB 76 TATTGCTGAGAGCAGACAGCGGGAGGAGTAAATGATCAGATATATAACCCAGGCATTTCGAGC 17
 QY 532 CGSACACGGCAACCC 546
 DB 16 TGGCAATGGCTACCC 2

RESULT 12
 US-09-949-016-13590/c
 ; Sequence 13590, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13590
 ; LENGTH: 149971
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(149971)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-13590

Query Match 57.8%; Score 367.2; DB 4; Length 149971;
 Best Local Similarity 80.9%; Pred. No. 1.4e-111;
 Matches 529; Conservative 0; Mismatches 83; Indels 42; Gaps 7;
 QY 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAACTA 60
 DB 18750 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAAGTA 18691
 QY 61 CAAATTTGTTCTTAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 DB 18690 TAAATGTTCTTAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 18640
 QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGCACCCTCCCGAG 180
 DB 18639 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGCACCCTCCCGAG 18580
 QY 181 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTT-AGAGCG 238
 DB 18579 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAAGAGCG 18520
 QY 239 GTCATCAGCAACCTCCCAACAGCACTGGGTTTCTGTTGAGAGGGGGGACTGAGAG 298

Db 18519 GTCGTGGCCAACTCCCAACAGCAGCTAGTTTCTCTGTTGAGCGGAGACTGAGAG 18460
Qy 299 ACAGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCTAGCTAGCTGGGAAGGTGA 358
Db 18459 ACAAGACTAGCTGGATTTCTTAGCGGACTAAGAATTCCTAAACCCAGCTGGGAAGGTGA 18400
Qy 359 CTGCATCCACTCTAAACATAGGGCTTGAACCTAGCTACACCCCGCCCAATC-----411
Db 18399 CGGCACCCACCTTTAAACACGGGGCTTGAACCTAGCTACACCCCAACCAATCAGATAGT 18340
Qy 412 --AGAGAGCTCACTAAATGCTAAATTAGGCAAAATAGGAGTAAAGAAATA-----G 462
Db 18339 AAGAGGGCTCACTAAATACAAATTAGCTAAGCAGGAGTAAAGAAATAGT 18280
Qy 463 CCAATCATCTATTGCTAGAGACACAGCGGGAGGACAAAGGATCGGGATATATAA-CCAG 521
Db 18279 CAAATCATATATACCTGAGAGCATAGGGGAGGACAAATGATGGGATATATAACCCCAA 18220
Qy 522 GCATTGAGCGGCAACGCCCTTTGGGTCCCTCCCTTTGTTGATGGCGCTCTGT 581
Db 18219 GCATCTTCCGGGAGTGGCAACCCCTTTGGGTCTCTCCCTCCATTTGATGGGCTCTGTT 18160
Qy 582 TTTCACTCTATTCACTCTATTAAATCTTGCAACTGMAAAAAAAGAAAAA 635
Db 18159 T-----CACTCTTAATCTTGCAATTGTGAGAAAAAAGAAAAA 18120

RESULT 13
US-09-949-016-15497/c
; Sequence 15497, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15497
; LENGTH: 245286
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(245286)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15497

Query Match 46.0%; Score 292; DB 4; Length 245286;
Best Local Similarity 92.5%; Pred. No. 4.2e-86;
Matches 307; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 288 GGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGC 347
Db 142457 GGTAGTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGCCTAGC 142398

Qy 348 TGGGAAGTGACTGCATCCACTCTAAACATGGGGCTTGCACCTTAGCTCACACCGGACC 407
Db 142397 TGGGAAGTGACTGCATCCACTCTAAACATGGGGCTTGCACCTTAGCTCACACCGGACC 407

Qy 408 AATCAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 467
Db 142338 TGGGAAGTGACTGCATCCACTCTAAACATGGGGCTTGCACCTTAGCTCACACCGGACC 142338

Qy 468 AATCAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 467
Db 142337 AATCAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 142278

Qy 468 CATCTATTGCTGAGAGACACAGCGGGAGGACAAGGATCGGGATATATAAACCCAGGCATTC 527

Db 142277 CATCTATCCCTGAGAGCACAGAGGGAGGACAATGATCAGGATATATAAACCCAGGCATTC 142218
Qy 528 GAGCGGCAACGCAACCCCTTTGGTCCCTCCCTTTGTTATGGGCGCTCTGTTTTCAC 587
Db 142217 CAGCGAGCAACGGCTACACTCTTTGGGTCCCTCCCTTTATATGGGAGCTCTGTTTTCAC 142158
Qy 588 TCTATTTCACTCTATTAAATCTTGCAACTGAA 619
Db 142157 TCTATTTCACTCTATTAAATCTTGCAACTGCA 142126

RESULT 14
US-09-949-016-133739/c
; Sequence 133739, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133739
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-133739

Query Match 45.9%; Score 291.6; DB 4; Length 601;
Best Local Similarity 92.2%; Pred. No. 1.2e-87;
Matches 306; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

Qy 288 GGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGC 347
Db 414 GGTAGTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGCCTAGC 355
Qy 348 TGGGAAGTGACTGCATCCACTCTAAACATGGGGCTTGCACCTTAGCTCACACCGGACC 407
Db 354 TGGGAAGTGACTGCATCCACTCTAAACATGGGGCTTGCACCTTAGCTCACACCGGACC 295
Qy 408 AATCAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 467
Db 294 AATCAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 235
Qy 468 CATCTATTGCTGAGAGCACAGCGGGAGGACAAGGATCGGGATATATAAACCCAGGCATTC 527
Db 234 CATCTATCCCTGAGAGCACAGAGGGAGGACAATGATCAGATATATAAACCCAGGCATTC 175
Qy 528 GAGCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGGCGCTCTGTTTTCAC 587
Db 174 CAGCGAGCAACGGCTACACTCTTTGGGTCCCTCCCTTTATATGGGAGCTCTGTTTTCAC 115
Qy 588 TCTATTTCACTCTATTAAATCTTGCAACTGAA 619
Db 114 TCTATTTCACTCTATTAAATCTTGCAACTGCA 83

RESULT 15
US-09-949-016-15393
; Sequence 15393, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED


```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 15393
; LENGTH: 194790
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15393

Query Match 45.4%; Score 288; DB 4; Length 194790;
Best Local Similarity 82.6%; Pred. No. 8.1e-85;
Matches 393; Conservative 0; Mismatches 70; Indels 13; Gaps 5;

QY 130 CTTGCTAGCCCTCTCCGATGTTAATGACATTAAGGACCCCTCCCGAGGAAATCTCA 189
DB 98345 CATAGTGGTCTTTTGGCAATTTCTGATGATCAAAAGGCATCCCTCTTGAGGAAATCTCA 98404

QY 190 ACTGCACAAACCCCTACTATGCCCAATTCAGCGGAAGCAATTAGAGCGGTCTACGCCA 249
DB 98405 ACTGCACAAACCCCTACTATGCCCAAGTTTCAGCAGGAAGCAAGTGTGTTGGCCA 98464

QY 250 ACCTCCCCCAACAGCAGCTTGGTTTCTGTTGAGAGGGGGGACTGAGAGACAGGACTAGC 309
DB 98465 ACCTCCCCCAACAGCAGCTTGGTTTCTGTTGAGAGGGGGGACTGAGAGACAGGACTAGC 98523

QY 310 TGGATTTCTTAGGCCCAACGAAATCCCTTAAGCTAGCT-GGGAAGGTGACTGCATCCAC 368
DB 98524 TGGATTTCTTAGGCCCAACGAAATCCCTTAAGCTAGCTGGGAAGGTGACCAACCCCTC 98583

QY 369 CTCTAAACATGGGGCTTGCACCTTAGCTCAGCTCACCCCGACCAATC-----AGAGAGCT 419
DB 98584 CTTTAAACACAGAGCTTGTAACTCAGCTCACCCCGACCAATC-----AGAGAGAGCT 98643

QY 420 CACTAAATATGCTAATTAGGC-AAAAATAGGAGGTAAAGAAAT-AGCCAATCATCTATTGC 477
DB 98644 CACTAAATATGCTAATTAGGC-AAAAATAGGAGGTAAAGAAAT-AGCCAATCATCTATTGC 98703

QY 478 CTGAGAGCAGACGGCGGAGGACAGGATCGGGATATAAACCCAGGCAATTCGAGCCGCCAA 537
DB 98704 CTGAGAGCAGACGGCGGAGGACAGGATCGGGATATAAACCCAGGCAATTCGAGCCGATC 98763

QY 538 CGGCAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATT 593
DB 98764 AGGTAACCCCTTTGGGTCCCTCCCTCCCTCAGCTGTATGGGAGCTCTGTTTTCACCTCTATT 98819

RESULT 16
US-09-949-016-44296/c
; Sequence 44296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 15086
; LENGTH: 219964
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15086

Query Match 42.4%; Score 269; DB 4; Length 219964;
Best Local Similarity 93.4%; Pred. No. 2.2e-76;
Matches 281; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

```

; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 44296
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-44296

Query Match 43.0%; Score 273.2; DB 4; Length 601;
Best Local Similarity 85.5%; Pred. No. 1.9e-81;
Matches 376; Conservative 1; Mismatches 49; Indels 14; Gaps 6;

QY 167 GCACCCCTCCGAGGAAATCTCAACTGACACACCCCTACTATGCCCCCAATTCAGCGGAA 226
DB 601 GCACCTCCTCTGAGGAAATCTCAACTGACACACCCCTACTATGCCCCCAATTCAGCGGAA 542

QY 227 GCAGTT-AGAGGGTTCATCAGCAACCTCCCAACAGCAGCTTGGGTTTTCCTCTTGAGAG 285
DB 541 GCAGTTAAGAGCGTTGTTGCGCAACTTCCCAACAGCAGCTTGGGTTTTCCTCTTGAGAT 482

QY 286 GGGGGAAGTCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAAACGAAATCCCTAAGCCTA 345
DB 481 GGGGGAAGTCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAAACGAAATTCCTAAGCCTA 422

QY 346 GGT-GGGAAGGTGACTGATCTCACTCTAAACATGGGGCTTGCACCTAGCTCAGCCCG 404
DB 421 GCTGGGGAAGGTGACCGCACCCACCTTTAAACACAGCGGCTGGTAACCTCAGCTCAGCCCTG 362

QY 405 ACCAATC-----AGAGAGCTCACTAAATGCTAATTTAGGC-AAAAATAGGAGGTAA 454
DB 361 ACCAATCGGGTAGTAAGAGGGCTCACTGAACACAAATTTAGGCTAAACGAGGAGGTAA 302

QY 455 AGAAATAG-CCAAATCATCTATTGCTGAGAGCAGCGGGAGGACAAAGATCGGGATAT 513
DB 301 RGAACAGTCAATCATATATGCTGAGAGCAGCAACAAG-GGGACAATGATCGGATAT 243

QY 514 AAACCCAGGATTCGAGCGCGCAACCGCAACCCCTTTGGGTCCCTCCCTTTGTATGGG 573
DB 242 AAACCTCAGGCAATTCAGAGCGGAGTGGCAACCCCTTTGGGTCCCTCCCTCGATGGG 183

QY 574 CGCTCTGTTTTCACCTCTATT 593
DB 182 AGCTCTGTTTTCACCTCTATT 163

```

```

RESULT 17
US-09-949-016-15086/c
; Sequence 15086, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 15086
; LENGTH: 219964
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15086

Query Match 42.4%; Score 269; DB 4; Length 219964;
Best Local Similarity 93.4%; Pred. No. 2.2e-76;
Matches 281; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

QY 293 TGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGA 352
Db 92554 TGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAATCCCTAAGCCTAGCTGGGA 92495
QY 353 AGGTGACTGCATCCACCTCTAAACATGCGGCTTGCAACTTACCTCACACCCGACCAATCA 412
Db 92494 AGGTGACCACGTCACCTTTAAAGATGCGGCTTGCAACTTACCTCACACCCGACCAATCA 92435
QY 413 GAGAGCTCACTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAAATAGCAATCATCT 472
Db 92434 GAGAGCTCACTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAAATAGCTAATCATCT 92375
QY 473 ATTGCTGAGACACAGCGGAGGACAAAGATCGGGATATAAACCCAGGCAATCGAGCC 532
Db 92374 ATCGCTTAAGACACAGCGGAGGACAAATGATCGGGATATAAACCCAGGCAATCGAGCC 92315
QY 533 GCGACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGGCGCTCTGTCTTCACTCTAT 592
Db 92314 GCGACGGCTACCTCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTCTTCACTCTAT 92255
QY 593 T 593
Db 92254 T 92254
RESULT 18
US-09-573-080A-385
; Sequence 385, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 385
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(780)
; OTHER INFORMATION: ltr17
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-385

Query Match 41.4%; Score 262.8; DB 4; Length 780;
Best Local Similarity 93.2%; Pred. No. 7.2e-78;
Matches 289; Conservative 0; Mismatches 12; Indels 9; Gaps 1;
QY 293 TGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGA 352
Db 1 TGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGA 60
QY 353 AGGTGACTGCATCCACCTCTAAACATGCGGCTTGCAACTTACCTCACACCCGACCAATC- 411
Db 61 AGGTGACCGCTTCCACCTTTAAACACGCGGCTTGCAACTTACCTCACACCCGACCAATCA 120
QY 412 -----AGAGAGCTCACTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAAATAGC 463

Db 121 GGTAGTAAAGAGAGCTCACTAAATGCTTAATTAGGCCAAAACAGGAGGTAAAGAAATAGC 180
QY 464 CAATCATCTATTTCCTTGAGAGCACAGCGGAGGAGCAAGGATCGGATATAAACCCAGGC 523
Db 181 CAATCATCTATTTCCTTGAGAGCACAGCGGAGGAGCAAGGATCGGATATAAACCCAGGC 240
QY 524 ATTTCGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGATGGCGCTCTGTGTT 583
Db 241 ATTTCGAGCGGCAACGCTACCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTGTT 300
QY 584 TCACTCTATT 593
Db 301 TCACTCTATT 310
RESULT 19
US-09-949-016-103638/c
; Sequence 103638, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103638
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-103638

Query Match 39.7%; Score 252.2; DB 4; Length 601;
Best Local Similarity 86.3%; Pred. No. 2.3e-74;
Matches 308; Conservative 1; Mismatches 29; Indels 19; Gaps 2;
QY 288 GGGACTGAGACACAGGACTAGCTGGATTTCCTAGGCCAACGAAATCCCTAAGCCTAGC 347
Db 472 GGTGATGAGACACAGGACTAGCTGGATTTCCTAGCTGACTAAGAATCCCTAAGCCTGCG 413
QY 348 TGGGAAGGTGACTGATCCACTCTAAACATGCGGCTTGCAACTTAGCTCACACCCGACC 407
Db 412 TGGGAAGGTGACCACTCACCTCTAAACACGCGGCTTGTAACCTTAGCTTACCCCGACC 353
QY 408 AATC-----AGAGAGCTCACTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAA 458
Db 352 AATCAGGTAGTAAAGAGAGCTCATTAATGCTTAATTAGGCCAAAACAGGARGTAAAGAA 293
QY 459 ATAGCAATCATCTATTTCCTGAGAGCACAGCGGAGGAGCAAGGATCGGATATAAAC 518
Db 292 GTAGCCAATCATCTGCTGCTGAGAGCACAGCGGAGGAGCAATGATCAGGATATAAAC 233
QY 519 GAGGATTCGAGCGCGCAACCGCAACCCCTTTGGGTCCCTCCCTTTGATGGCGCTC 578
Db 232 CAGTCATTCAGCTGCGCAACCGCTACCCACTTTGGGTCTCTCCCTTTGATGGAGCTC 173
QY 579 TGTCTTCACTCTATTTCCTACTCTATTAAATCTTTGCAACTGAAAAAAA 635
Db 172 TGT-----TTTCACTCTATTAAATCTTTGCAACTGCAAAAAA 126

RESULT 20
US-09-949-016-103639/c
; Sequence 103639, Application US/09949016

QY 408 AATC-----AGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAA 458
|||
Db 116297 AATCAGGTAGTAAGAGAGCTCATTAATAATGCTAATTAGGAAAACAGGAAGTAAAGAA 116238
|||
QY 459 ATAGCAATCATCTATTGCTCGAGAGCACAGCGGAGGACAAAGGATCGGATATAAAACC 518
|||
Db 116237 GTAGCCATCATCTGCTCGCTGAGAGCACAGCGGAGGAGCAATGATCAGGATATAAAACC 116178
|||
QY 519 CAGGCATTGAGCGCGCAACGCAACCCCTTTGGGTCCCTTCTTTGTTGATGGCGGCTC 578
|||
Db 116177 CAGTCATTCAGTGCACAGCGCTACCCACATTTGGGTCTCTCTCCCTTTGTTGATGGAGCTC 116118
|||
QY 579 TCTTTTCACTCTATTTCACCTCTATTAAATCTTGCAACTGAAAAAAGAAAAA 635
|||
Db 116117 TGT-----TTTCACTCTATTAAATCTTGCAACTGCAAAAAAAGAAAAA 116071
|||

RESULT 23

US-09-949-016-44299/c
; Sequence 44299, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44299
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44299

Query Match 39.3%; Score 249.4; DB 4; Length 601;
Best Local Similarity 88.5%; Pred. No. 2e-73;
Matches 284; Conservative 1; Mismatches 27; Indels 9; Gaps 1;

QY 282 AGAGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAACGAAATCCCTAAG 341
|||
Db 601 AGAAGGAGACTGAGAGACAGGACTGACTGGATTTCTTAGGCCGACTAAGAATTCCTAAG 542
|||
QY 342 CTTAGCTGGGAAGGTGACTGCATCCACCTTAACATGCGGCTTGCAACTAGCTCACAC 401
|||
Db 541 CTTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGCGGCTTGCAACTAATCATAT 482
|||
QY 402 CGACCAATC-----AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGT 452
|||
Db 481 CTGACCAATCAGGTAGTAAGAGAGTTCATAAATGCTAATAGGCAAAACAGAGGT 422
|||
QY 453 AAGAAATAGCCAAATCATCTATTGCTGAGAGCACAGCGGAGGGAACAAGGATCGGGATA 512
|||
Db 421 AAGAAATAGCCAAATCATCTATTGCTGAGAGCACAGCGGAGGGAATAATGATCAGGATA 362
|||
QY 513 TAAACCGAGCATTCAGCGCGCAACGCCCTTTGGGTCCCTCCCTCTTGTATGG 572
|||
Db 361 TAAACCGAGCATTTGAGCTGGCAATGGCTACCCCTCTTTGGGTCCCTCTTGTATGG 302
|||
QY 573 GCCTCTGTTTTCACCTCTATT 593
:|||||
Db 301 RAGCTCTGTTTTCACCTCTATT 281
:|||||

RESULT 24

US-09-949-016-17420

; Sequence 17420, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17420
; LENGTH: 84571
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84571)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17420

Query Match 38.7%; Score 245.8; DB 4; Length 84571;
Best Local Similarity 89.4%; Pred. No. 8e-71; Mismatches 0; Indels 14; Gaps 3;
Matches 303; Conservative 0;

QY 291 ACTGAGACAGGACTAGCTGGATTTCTTAGGCCAAACGAA-GAATCCCTAAGCCTTAGCTG 349
|||
Db 7322 ACTGAGACAGGACTAGCTGGATTTCTTAGGCCGACTAAGGAATCCCAAGCCTTAGCTG 7381
|||
QY 350 GGAAGGTGACTGCATCCACCTTAACATGGGGTTGCAACTAGCTCACACCCGACCAA 409
|||
Db 7382 GGAAGTGACCGCTTCCACCTTTAAACATGGGACTTGCAACTAGCTCACACCCGACCAA 7441
|||
QY 410 TCA-----GAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAAT 460
|||
Db 7442 TCAGATAGTAAGGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGAAAT 7501
|||
QY 461 AGCCAAATCATCTATTGCTGAGAGCACAGCGGAGGGAACAAGGATCGGGATATAAACCCA 520
|||
Db 7502 AGC-----CATCTGCTGCTGAGAGCACAGGAGGAGCAATGATCAGGATATAAACCCA 7557
|||
QY 521 GGCATTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTGTATGGCGGCTCTG 580
|||
Db 7558 GGCATTCGAGCGGCAATGGCAACCCCTTTGGGTCCCTCCCTTGTACGGGAGCTCTG 7617
|||
QY 581 TTTTCACTCTATTTCACCTCTAATAATCTTGCAACTGAA 619
|||
Db 7618 TTTTCACTGTGTTTCACCTCTAATAATCTTGCAACTGCA 7656
|||

RESULT 25

US-09-949-016-17002
; Sequence 17002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 17068
 LENGTH: 140224
 TYPE: DNA
 ORGANISM: Human
 NAME/KEY: misc feature
 LOCATION: (1)...(140224)
 OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17068

Query Match 38.7%; Score 245.6; DB 4; Length 140224;
 Best Local Similarity 86.9%; Pred. No. 1.3e-70;
 Matches 285; Conservative 0; Mismatches 34; Indels 9; Gaps 1;
 QY 279 TTGAGAGGGGGAGTGTAGAGACGAGGACTAGCTGGATTTCCTAGGCCAACGAAGATCCCT 338
 Db 16074 TAGCAGGCCCTTATTCAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCT 16133
 QY 339 AAGCTAGCTGGGAGGTGACTGCATCCACTCTTAACATGGGGCTTGCAACTTAGCTCA 398
 Db 16134 AAGCTAGCTGGGAGGTGACTGCATCCACTCTTAACATGGGGCTTGCAACTTAGCTCA 16193
 QY 399 CACCGACCAATC-----AGAGAGCTCACTAAATGCTAATAGGCAAAATAGGA 449
 Db 16194 CACCGACCAATCAGGTAGTAAGAGAGCTCACTAAATGCTAATAGGCAAAATAGGA 16253
 QY 450 GGTAAGAAATAGCAATCATCTATTCCTGTAGAGCACAGCGGAGGGAACAAGATCGGG 509
 Db 16254 GGGAAAGAAATAGCAATCATCTATTCCTGTAGAGCACAGCGGAGGGAACAATGTCAGG 16313
 QY 510 ATATAAACCCAGGATTCGAGCGGGAACGGCAACCCCTTGGTCCCTCCCTTTGTA 569
 Db 16314 ATATAAACCCAGGATTCGAGCGGGAACGGCAACCCCTTGGTCCCTCCCTTTGTA 16373
 QY 570 TGGCGCTCTGTTTCACTCTATTTCAC 597
 Db 16374 TGGAGCTCCGTTTCACTCTATTTCAC 16401

RESULT 26
 US-09-949-016-17068/c
 Sequence 17068, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 17068
 LENGTH: 89584
 TYPE: DNA
 ORGANISM: Human
 US-09-949-016-17068

Query Match 37.7%; Score 239.4; DB 4; Length 89584;
 Best Local Similarity 89.8%; Pred. No. 1.2e-68;
 Matches 283; Conservative 0; Mismatches 21; Indels 11; Gaps 2;
 QY 288 GGGACTGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAGATCCCTTAAGCCTAGC 347
 Db 66011 GGAATTAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAGATCCCTTAAGCCTAGC 65952

QY 348 TGGAGAGGTGACTGCATCCACCTCTAAACATCGGGCTTCGAACCTTAGCTCACACCCGACC 407
 Db 65951 TGGAGAGGTGATCACATCCACCTTTAAACATCGGGCTTACAACCTTAGCTCACACCCGACC 65892
 QY 408 AATCAG-----AGACTCACTAAATGCTAATAGGCAAAATAGGCTTAAGGA 458
 Db 65891 AATCAGGTAGTAAGAGAGCTCACTAAATGCTAATAGGCAATAA--AGAGGTAAGAA 65834
 QY 459 ATAGCCAATCATCTATTGCTGTAGAGCACAGCGGAGGACAAAGGATCGGGATATAAAC 518
 Db 65833 ATAGCCAATCATCTATTGCTGTAGAGCACAGCGGAGGACAAATGATCGGGATATAAAC 65774
 QY 519 CAGGCAATTCAGCGCGCAACCGCAACCCCTTTGGGTCCCTCCCTTTGATGGGCGCTC 578
 Db 65773 CAGGCAATTCAGCGCGCAATGCTACCCCTTTGGGTCCCTCTCTTTGATGGGAGCTC 65714
 QY 579 TGTTTCACTCTATT 593
 Db 65713 TGTTTCACTCTATT 65699

RESULT 27
 US-08-686-878A-50
 Sequence 50, Application US/08686878A
 Patent No. 5708157
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 APPLICANT: McCoy, John
 APPLICANT: Lavallie, Edward
 APPLICANT: Racie, Lisa
 APPLICANT: Merberg, David
 APPLICANT: Treacy, Maurice
 APPLICANT: Evans, Cheryl
 APPLICANT: Spaulding, Vikki
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 NUMBER OF SEQUENCES: 71
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genetics Institute, Inc.
 STREET: 87 CambridgePark Drive
 CITY: Cambridge
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/686,878A
 FILING DATE:
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Brown, Scott A.
 REGISTRATION NUMBER: 32,724
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 279 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-686-878A-50

Query Match 36.8%; Score 233.6; DB 1; Length 279;
 Best Local Similarity 88.1%; Pred. No. 2.6e-68;
 Matches 245; Conservative 1; Mismatches 32; Indels 0; Gaps 0;

Query Match 36.8%; Score 233.6; DB 1; Length 279;
Best Local Similarity 88.1%; Pred. No. 2.6e-68;
Matches 245; Conservative 1; Mismatches 32; Indels 0; Gaps 0

QY 358 ACTGTCACCTTAACATGGGGCTTGCAACTTAGCTACACCCGACCAATCAGAG 417

Query Match	34.6%	Score	219.6	DB 4	Length	8523			
Best Local Similarity	87.7%	Pred. No.	1.2e-62						
Matches	256	Conservative	0	Mismatches	24	Indels	12	Gaps	1
Qy	1	CCCTGTATCTTTAACTTCCTTGGTTAAAGTTGTCTCTCTCCAGAACTCAAAACGTGTAAAACTA	60						
Db	8244	CCCTGTATCTTTAACTTCCTTGGTTAACTTGTCTCTCTCCAGAATCGAAGCTGTAAAACTA	8303						
Qy	61	CAAAATGTTCTTTCAAAATGGAGACCAAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120						
Db	8304	-----CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTTACCGCAGACCC	8351						
Qy	121	CTGGACCGGCCTGCTAGCCCAATGCTCCGATGTTAAATGACATTTGAAGGCACCCCTCCCGAG	180						
Db	8352	CTGGACCGGCCTGCTAGCCCAAGATCTGATGTTAATGACATCAAGGCACCCCTCCTTGAG	8411						
Qy	181	GAAATCTCAACTGCACAAACCCCTACTATGTCCCAATTCAGCGGGAAGCAGTTAGACGGGT	240						
Db	8412	GAAATCTCAGCTGCACAACTCTACTAGCCCAATTCAGCAGGAAGCAGTTAGACGGGT	8471						

```
QY 241 CATCAGCCACCTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGAC 292
Db 8472 CGTCGGCCAACTCCCAACAGCAGCACTAGATTCTCTGTTGAGATGGGGAC 8523

RESULT 30
US-09-621-976-9366/C
; Sequence 9366, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9366
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9366

Query Match 31.0%; Score 196.6; DB 4; Length 456;
Best Local Similarity 84.9%; Pred. No. 1.1e-55;
Matches 264; Conservative 0; Mismatches 24; Indels 23; Gaps 3;

QY 292 CTGAGAGACAGGACTAGCTGGATTCTCTAGGCCAAACGAAGAATCCCTTAAGCCCTAGCTGGG 351
Db 307 CTGAGAGACAGGACTAGCTGGATTCTCTAGGCCGACTAGAAATCCCTTAAGCCCTAGCTGGG 248

QY 352 AAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCCAATC 411
Db 247 AAGGTGACCACTCCACCTTTAAGCAGGGGGCTTGCAACTTAGCTCACACCCGACGTC 188

QY 412 -----AGAGACTCATAAATGCTAAATAGGCAAAAATAGGAGTAAAGAAATAG 462
Db 187 AGGTAGTAAAGCGAGCTCACTAAATGCTAAATAGGCTAAACAGGAGGTAAAC-CAATAG 129

QY 463 CCAATCATCTATTGCTCTGAGAGACAGCGGGGAGCAAGGATCGGGATATAACCCAGG 522
Db 128 CCAATCATCTATTCTCTGAGAGACAGCGGGGAGCAATGATAGGATATAAC----- 74

QY 523 CATTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCTCTGTTATGGCGCTCTGTT 582
Db 73 -----CTGGCAACGGCTCTCTGGGTCCCTCTCTGTTATGGGAGCTCTGTT 22

QY 583 TTCACTCTATT 593
Db 21 TTCACTCTATT 11

RESULT 31
US-09-949-016-128750
; Sequence 128750, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17057
; LENGTH: 154023
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-128750

Query Match 28.4%; Score 180.4; DB 4; Length 154023;
Best Local Similarity 82.1%; Pred. No. 1.3e-48;
Matches 271; Conservative 0; Mismatches 46; Indels 13; Gaps 5;

QY 293 TGAGAGACAGGACTAGCTGGATTCTCTAGGCCAAACGAAGAATCCCTTAAGCCTAGCT-GGG 351
Db 92234 TGAGAGACAGGACTAGCTGGATTCTCTAGGCTGACTAAGAAATTCCTAAGTCTATCTGGGG 92293

QY 352 AAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCCAATC 411
Db 92294 AAGGTGACTGCAGCCCACTTTAAACATGGGGCTTGTAACCTAGCTCACCTGACCAATC 92353
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128750
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-128750

Query Match 30.8%; Score 195.6; DB 4; Length 601;
Best Local Similarity 80.5%; Pred. No. 2.7e-55;
Matches 280; Conservative 1; Mismatches 55; Indels 12; Gaps 4;

QY 130 CTGCTAGCCCATGCTCCGATGTTTATGACATTGAAGGCACCCCTCCCGAGGAATCTCA 189
Db 242 CATAGTGGTCTCTTTGGCCATTCTGATGACATCAAGGCATCCCTCTTGAGGAATCTCR 301

QY 190 ACTGCACACACCCCTACTATGCCCAATTCAGGCGGAAGCAGTTAGAGCGGTCAATCAGCCA 249
Db 302 ACTGCACACACCCCTACTATGCCCAATTCAGGCGGAAGCAGTTAGAGCGGTCAATCAGCCA 361

QY 250 ACTCTCCCAACAGCAGCTTTGGGTTTCTCTGTTGAGAGGGGGGACTGAGAGACAGGACTAGC 309
Db 362 ACTCTCCCAACAGCAGCTTTGGGTTTCTCTGTTGAGAGGGGGGACTGAGAGACAGGACTAGC 420

QY 310 TGGATTTCCTAGGCCAAACGAAGAATCCCTTAAGCCTAGCT-GGGAAGGTGACTGCATCCAC 368
Db 421 TGGATTTCCTAGGCCGAGCTTAAGAAATTCCTTAAGCCTAGCTGCACACCCGAGGTACCAACCCCTC 480

QY 369 CTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC-----AGAGAGCT 419
Db 481 CTCTAAACATGGGGCTTGCAACTTAGCTGCACACCCGAGGTACCAACCCGAGGTACCAACCCCTC 540

QY 420 CACTAAATGCTAAATAGGC-AAAAATAGGAGGTAAAGAAATAGCCAA 466
Db 541 CACTAAATGCTAAATAGGCTTAAACACAGGAGTAAAGAAATAGCCAA 588
```

```
RESULT 32
US-09-949-016-17057
; Sequence 17057, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17057
; LENGTH: 154023
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17057

Query Match 28.4%; Score 180.4; DB 4; Length 154023;
Best Local Similarity 82.1%; Pred. No. 1.3e-48;
Matches 271; Conservative 0; Mismatches 46; Indels 13; Gaps 5;

QY 293 TGAGAGACAGGACTAGCTGGATTCTCTAGGCCAAACGAAGAATCCCTTAAGCCTAGCT-GGG 351
Db 92234 TGAGAGACAGGACTAGCTGGATTCTCTAGGCTGACTAAGAAATTCCTAAGTCTATCTGGGG 92293

QY 352 AAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCCAATC 411
Db 92294 AAGGTGACTGCAGCCCACTTTAAACATGGGGCTTGTAACCTAGCTCACCTGACCAATC 92353
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 195792
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-195792

Query Match 27.8%; Score 176.8; DB 4; Length 601;
Best Local Similarity 85.4%; Pred. No. 6e-49;
Matches 270; Conservative 0; Mismatches 32; Indels 14; Gaps 6;

QY 292 CTGAGAGACAGGACTAGCTGGATTTCCTAGGCGCAACGAAGATCCCTTAAGCCTAGCT-GG 350
Db 540 CTGAGAGACAGGACTAGCTGGATTTCCTAGGCGCTAGGAAATCTTTAAGCCTAGCTGGG 481
QY 351 GAAGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCAGTCCACCCGACCAAT 410
Db 480 GAAGTGACCGCACCCACCTTTTAAACACGGAGCTTGTAACTCAGTCCACCCGACCAAT 421
QY 411 C-----AGAGAGCTCACTAAATGCTAAATAGGC-AAAAATAGAGGTAAGAAAT 460
Db 420 CAGGTAGTAAAGGGGCTCACTAAATACAAATTTAGGCTAAGACGAGGTAAGAAAT 361
QY 461 AG-CCAATCATCTATTGCTGAGAGCAGCGGGGAGGAGCAAGGATCGGATATAAACCC- 518
Db 360 AGTCAAAATCATCTATCATCTGAGAGCAGCGGGGAGGAGCAAGTATGGGATATAAACCCY 301
QY 519 CAGGCATTTCGAGCGGCA-ACGGCAACGCCCTTTGGGTCCTCCCTCTTTGTATGGGCGCT 577
Db 300 CAGGCATTTCGAGCGGAGTGGGCAACCCCTTTGGGTCCTCCCTCTTTGTATGGGAGCT 241
QY 578 CTGTTTTCACCTCTATT 593
Db 240 CTGTTTTCACCTCTATT 225

RESULT 35
US-09-949-016-17296
; Sequence 17296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 17296
; LENGTH: 251672
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 195792
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-195792

Query Match 28.0%; Score 178; DB 4; Length 601;
Best Local Similarity 88.5%; Pred. No. 2.4e-49;
Matches 216; Conservative 0; Mismatches 25; Indels 3; Gaps 2;

QY 1 CCCTGTATCTTTAACTCCCTGTTTAAAGTTTGTCTCTTCCAGAAATCAAAATCTGTAAGCTA 60
Db 242 CCCTGTATCTTTAACTCCCTGTTTAAAGTTTGTCTCTTCCAGAAATCAAAATCTGTAAGCTA 183
QY 61 CAAATGTTCTTCAATGGAGCAGCAGATGGAGTCCATGATCAAGATCCACCGTGACCC 120
Db 182 CAAATGTTCTTCAATGGAGCAGCAGATGGAGTCCATGATCAAGATCCACCGTGACCC 123
QY 121 CTGACCGGCTGTACGCCATGCTCCGATGTTTAAATGACATTCGAAGCAGCCCTCCCGAG 180
Db 122 --GACGAGCTGTACGCCATGCTCC-AGTTTAAATCAATCAAGGCACTCTCCAGAG 66
QY 181 GAAATCTCAATGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTATAGAGCGGT 240
Db 65 GAAATCTCAATGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTATAGAGCAAT 6
QY 241 CATC 244
Db 5 CATC 2

RESULT 14
US-09-949-016-195792/c
; Sequence 195792, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 44301
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-44301

Query Match 28.0%; Score 178; DB 4; Length 601;
Best Local Similarity 88.5%; Pred. No. 2.4e-49;
Matches 216; Conservative 0; Mismatches 25; Indels 3; Gaps 2;

QY 1 CCCTGTATCTTTAACTCCCTGTTTAAAGTTTGTCTCTTCCAGAAATCAAAATCTGTAAGCTA 60
Db 242 CCCTGTATCTTTAACTCCCTGTTTAAAGTTTGTCTCTTCCAGAAATCAAAATCTGTAAGCTA 183
QY 61 CAAATGTTCTTCAATGGAGCAGCAGATGGAGTCCATGATCAAGATCCACCGTGACCC 120
Db 182 CAAATGTTCTTCAATGGAGCAGCAGATGGAGTCCATGATCAAGATCCACCGTGACCC 123
QY 121 CTGACCGGCTGTACGCCATGCTCCGATGTTTAAATGACATTCGAAGCAGCCCTCCCGAG 180
Db 122 --GACGAGCTGTACGCCATGCTCC-AGTTTAAATCAATCAAGGCACTCTCCAGAG 66
QY 181 GAAATCTCAATGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTATAGAGCGGT 240
Db 65 GAAATCTCAATGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTATAGAGCAAT 6
QY 241 CATC 244
Db 5 CATC 2

RESULT 33
US-09-949-016-44301/c
; Sequence 44301, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 44301
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-44301

Query Match 27.8%; Score 176.8; DB 4; Length 601;
Best Local Similarity 85.4%; Pred. No. 6e-49;
Matches 270; Conservative 0; Mismatches 32; Indels 14; Gaps 6;

QY 292 CTGAGAGACAGGACTAGCTGGATTTCCTAGGCGCAACGAAGATCCCTTAAGCCTAGCT-GG 350
Db 540 CTGAGAGACAGGACTAGCTGGATTTCCTAGGCGCTAGGAAATCTTTAAGCCTAGCTGGG 481
QY 351 GAAGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCAGTCCACCCGACCAAT 410
Db 480 GAAGTGACCGCACCCACCTTTTAAACACGGAGCTTGTAACTCAGTCCACCCGACCAAT 421
QY 411 C-----AGAGAGCTCACTAAATGCTAAATAGGC-AAAAATAGAGGTAAGAAAT 460
Db 420 CAGGTAGTAAAGGGGCTCACTAAATACAAATTTAGGCTAAGACGAGGTAAGAAAT 361
QY 461 AG-CCAATCATCTATTGCTGAGAGCAGCGGGGAGGAGCAAGGATCGGATATAAACCC- 518
Db 360 AGTCAAAATCATCTATCATCTGAGAGCAGCGGGGAGGAGCAAGTATGGGATATAAACCCY 301
QY 519 CAGGCATTTCGAGCGGCA-ACGGCAACGCCCTTTGGGTCCTCCCTCTTTGTATGGGCGCT 577
Db 300 CAGGCATTTCGAGCGGAGTGGGCAACCCCTTTGGGTCCTCCCTCTTTGTATGGGAGCT 241
QY 578 CTGTTTTCACCTCTATT 593
Db 240 CTGTTTTCACCTCTATT 225

RESULT 35
US-09-949-016-17296
; Sequence 17296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 17296
; LENGTH: 251672
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
```



```
; LOCATION: (1)...(251672)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17296

Query Match      27.8%; Score 176.8; DB 4; Length 251672;
Best Local Similarity 85.4%; Pred. No. 3e-47;
Matches 270; Conservative 0; Mismatches 32; Indels 14; Gaps 6;

QY 292 CTGAGACAGGACTAGCTGGATTCTTAGGCCAACGAGAAATCCCTAAGCCTAGCT-GG 350
DB 42623 CTGAGACAGGACTAGCTGGATTCTTAGGCTAGTGAATCTTAAAGCCTAGCTGG 42682

QY 351 GAAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAAT 410
DB 42683 GAAGGTGACCGCACCTTTAAACACGGAGCTTGTAACTCAGCTCACACCCCAACCAAT 42742

QY 411 C-----AGAGAGCTCACTAAATGCTAAATTAGGC-AAAAATAGGAGGTAAAGAAAT 460
DB 42743 CAGGTAGTAAAGAGGGCTCACTAAATACAAATTAGGCTAAGAGCAGGAGGTAAAGAAAT 42802

QY 461 AG-CCAAATCATCTATTCCCTGAGACACAGCGGGGAGGACAAAGATCGGGATATAAA-CC 518
DB 42803 AGTCAAATCATCTATCTATCTGAGACACAGCGGGGAGGACAAATGATTTGGGATATAAACCC 42862

QY 519 CAGGCATTGAGCGCGCA-ACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCT 577
DB 42863 CAGGCATTGAGCGGGAGTGGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCT 42922

QY 578 CTGTTTTCACTCTATT 593
DB 42923 CTGTTTTCACTCTATT 42938

RESULT 36
US-09-949-016-11973
; Sequence 11973, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11973
; LENGTH: 251682
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc.feature
; LOCATION: (1)...(251682)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11973

Query Match      27.8%; Score 176.8; DB 4; Length 251682;
Best Local Similarity 85.4%; Pred. No. 3e-47;
Matches 270; Conservative 0; Mismatches 32; Indels 14; Gaps 6;

QY 292 CTGAGACAGGACTAGCTGGATTCTTAGGCCAACGAGAAATCCCTAAGCCTAGCT-GG 350
DB 42623 CTGAGACAGGACTAGCTGGATTCTTAGGCTAGTGAATCTTAAAGCCTAGCTGG 42692

QY 351 GAAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAAT 410
DB 42693 GAAGGTGACCGCACCTTTAAACACGGAGCTTGTAACTCAGCTCACACCCCAACCAAT 42752
```

```
QY 411 C-----AGAGAGCTCACTAAATGCTAAATTAGGC-AAAAATAGGAGGTAAAGAAAT 460
DB 42753 CAGGTAGTAAAGAGGGCTCACTAAATACAAATTAGGCTAAGAGCAGGAGGTAAAGAAAT 42812

QY 461 AG-CCAAATCATCTATTCCCTGAGACACAGCGGGGAGGACAAAGATCGGGATATAAA-CC 518
DB 42813 AGTCAAATCATCTATCTATCTGAGACACAGCGGGGAGGACAAATGATTTGGGATATAAACCC 42872

QY 519 CAGGCATTGAGCGCGCA-ACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCT 577
DB 42873 CAGGCATTGAGCGGGAGTGGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCT 42932

QY 578 CTGTTTTCACTCTATT 593
DB 42933 CTGTTTTCACTCTATT 42948

RESULT 37
US-09-949-016-44310/c
; Sequence 44310, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44310
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-44310

Query Match      27.6%; Score 175; DB 4; Length 601;
Best Local Similarity 82.6%; Pred. No. 2.4e-48;
Matches 262; Conservative 1; Mismatches 41; Indels 13; Gaps 5;

QY 289 GGACTGAGACAGGACTAGCTGGATTCTTAGGCCAACGAGAAATCCCTAAGCCTAGCT 348
DB 521 GTATTGAGACAGGACTAGCTGGATTCTTAGGTTGACTTAAGAAATTCCTAAGCCTAGCT 462

QY 349 -GGGAGGTGACTGCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACC 407
DB 461 GGGGAGGTGACCACACCCACCTTTAAACACAGCGCTGGTAACTCAGCTCACACTGACC 402

QY 408 AATC-----AGAGAGCTCACTAAATGCTAAATTAGGC-AAAAATAGGAGGTAAAGA 457
DB 401 AATCAGGTAGTAAAGAGGCTCACTAAATACAAATTAGGCTAAGAGCAGGAGGTAAAGA 342

QY 458 AATAG-CCAAATCATCTATTCCCTGAGAGCAAGCGGGAGGACAAAGATCGGGATATAAA 516
DB 341 AATAGTCAAATCATATATACCTGAGAGCAACACGGG-GGGWCAATGATCAGGATATAAA 283

QY 517 CCAGGCAATTCGAGCGGCAACGCCACCCCTTTGGGTCCCTCCCTTTGTATGGGCGC 576
DB 282 CCTAGGCATTGAGCAGGAGTGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGC 223

QY 577 TCTGTTTTCACTCTATT 593
DB 222 TCTGTTTTCACTCTATT 206

RESULT 38
```

```
US-09-949-016-13413/c
; Sequence 13413, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13413
; LENGTH: 116652
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(116652)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13413

Query Match      27.5%; Score 174.6; DB 4; Length 116652;
Best Local Similarity 84.9%; Pred. No. 1e-46;
Matches 269; Conservative 0; Mismatches 34; Indels 14; Gaps 6;

QY 291 ACTGAGAGACGAGCTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCTAGCT-G 349
DB 67040 AGTGAGAGACGAGCTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCTAGCTGG 66981

QY 350 GGAAGTGACTGATCCTCTAAACATGGGGCTTGGCACTTAGCTCAGCCGACCAA 409
DB 66980 GGAAGGTGACCGCACCCACCTTTAAACAAGGGGCTTGGCACTTAGCTCAGCCGACCAA 66921

QY 410 TC-----AGAGAGCTCACTAAATGCTAATTAGGC-AAAAATAGGAGGTAAAGAAA 459
DB 66920 TCAGGTAGTAAAGAGGGCTCACTAAATACCAATTAGGCTAAAGCAGGAGGTAAAGAAA 66861

QY 460 TAG-CCAATCATCTATTGCTGAGACACAGCGGAGGAGCAAGGATCGGGATATAA-C 517
DB 66860 TAGTCAATAATCTATCATCTGAGACACAGGGGAGGAGCAATGATGGGATATAAACCC 66801

QY 518 CAGGCAATTCAGCCGCA-ACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGC 576
DB 66800 CAGGCAATTCAGCTGGGAGTGGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGC 66741

QY 577 TCTGTTTCACTCTATT 593
DB 66740 TCTGTTTCACTCTATT 66724

RESULT 39
US-09-949-016-13633/c
; Sequence 13633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13633
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13633

Query Match      27.4%; Score 174; DB 4; Length 39686;
Best Local Similarity 83.1%; Pred. No. 7.9e-47;
Matches 261; Conservative 0; Mismatches 40; Indels 13; Gaps 5;

QY 293 TGAGAGACGAGCTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCTAGCT-GGG 351
DB 6761 TGAGAGCAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCTAGCTGGGG 6702

QY 352 AAGTGACTGATCCACCTCTAAACATGGGGCTTGGCACTTAGCTCAGCCGACCAATC 411
DB 6701 AAGTGACTGATCCACCTCTTAAACACGGGGCTTGGCACTTAGCTCAGCTCAGCTCAGCC 6642

QY 412 -----AGAGAGCTCACTAAATGCTAATTAGG-CAAAAAATAGGAGGTAAAGAAATA 461
DB 6641 AGTGTTGTTAAAGAGGCTCACTAAATATCAATTAGGTTTAAAGCAGGAGGTAAAGAAATA 6582

QY 462 -GCCAATCATCTATTGCTGAGACACAGCGGAGGAGCAAGGATCGGGATATAA-CCC 519
DB 6581 CTCAAAATCATCTATCTGAGAGCAGCGGAGGAGCAAGGATCGGGATATAAACC 6522

QY 520 AGGCAATTCAGCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCT 579
DB 6521 AGGCAATTCAGCGGCAACGCGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCT 6462

QY 580 GTTTTCACTCTATT 593
DB 6461 GTTTTCACTCTATT 6448

RESULT 40
US-09-949-016-11770/c
; Sequence 11770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11770
; LENGTH: 49487
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(49487)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11770

Query Match      27.4%; Score 174; DB 4; Length 49487;
Best Local Similarity 83.1%; Pred. No. 9.1e-47;
```

```
US-09-949-016-13633/c
; Sequence 13633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13633
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13633

Query Match      27.4%; Score 174; DB 4; Length 39686;
Best Local Similarity 83.1%; Pred. No. 7.9e-47;
Matches 261; Conservative 0; Mismatches 40; Indels 13; Gaps 5;

QY 293 TGAGAGACGAGCTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCTAGCT-GGG 351
DB 6761 TGAGAGCAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCTAGCTGGGG 6702

QY 352 AAGTGACTGATCCACCTCTAAACATGGGGCTTGGCACTTAGCTCAGCCGACCAATC 411
DB 6701 AAGTGACTGATCCACCTCTTAAACACGGGGCTTGGCACTTAGCTCAGCTCAGCTCAGCC 6642

QY 412 -----AGAGAGCTCACTAAATGCTAATTAGG-CAAAAAATAGGAGGTAAAGAAATA 461
DB 6641 AGTGTTGTTAAAGAGGCTCACTAAATATCAATTAGGTTTAAAGCAGGAGGTAAAGAAATA 6582

QY 462 -GCCAATCATCTATTGCTGAGACACAGCGGAGGAGCAAGGATCGGGATATAA-CCC 519
DB 6581 CTCAAAATCATCTATCTGAGAGCAGCGGAGGAGCAAGGATCGGGATATAAACC 6522

QY 520 AGGCAATTCAGCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCT 579
DB 6521 AGGCAATTCAGCGGCAACGCGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCT 6462

QY 580 GTTTTCACTCTATT 593
DB 6461 GTTTTCACTCTATT 6448

RESULT 40
US-09-949-016-11770/c
; Sequence 11770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11770
; LENGTH: 49487
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(49487)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11770

Query Match      27.4%; Score 174; DB 4; Length 49487;
Best Local Similarity 83.1%; Pred. No. 9.1e-47;
```

	Matches	261;	Conservative	0;	Mismatches	40;	Indels	13;	Gaps	5;
Qy	293	TGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAGAAATCCCTAAGCCTAGCT-GGG	351							
Db	16051	TGAGAGGCGAGGACTAGCTGGATTTCCTAGGCCGACTAAGAAATTCCTAAGCCTAGCTGGGG	15992							
Qy	352	AAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTTAGCTCACACCCGACCAATC	411							
Db	15991	AAGGTGACTGCACCCACCTTTAAACACGGGGCTTGTAACTCAGCTCACACTCGACCAATC	15932							
Qy	412	-----AGAGAGCTCACTAAATGCTTAATTAGG-CAAAATAGGAGGTAAAGAAATA	461							
Db	15931	AGGTGGTAAAGAAGGCTCACTAAATATCAATTAGGTTAAAGCAGGAGGTAAAGAAATA	15872							
Qy	462	-GCCAATCATCTATTGCCTGAGAGCACAGCGGGGAGGACAGGATCGGATATAAA-CCC	519							
Db	15871	CTCAAAATCATCTATCATCTGAGAGCACAGGGGGAGGACAAATGATGGGATATAAACCCC	15812							
Qy	520	AGGCATTGAGCGGGCAACCGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCT	579							
Db	15811	AGGCATTGAGCAAGGAGGGGCAACCCCTTTGGATCCCTTACCTTTGTATAGCAGCTCT	15752							
Qy	580	GTTTTCACCTATT	593							
Db	15751	GTTTTCACCTATT	15738							

Search completed: February 21, 2005, 15:01:18
 Job time : 120.885 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 11:52:50 ; Search time 350.955 Seconds
(without alignments)
10694.149 Million cell updates/sec

Title: US-09-319-156B-6

Perfect score: 635

Sequence: 1 cccgtatcttttaacctctc.....tgaaaaaaaaaaaaaaaaaa 635

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	635	100.0	635	8	US-08-979-847-102
2	635	100.0	635	16	US-10-114-104-102
3	617.4	97.2	2030	17	US-10-637-793-25
4	564.8	88.9	366710	18	US-10-719-993-6792
5	546.4	86.0	2074	17	US-10-416-642-4
6	545	85.8	1329	8	US-08-979-847-108
7	545	85.8	1329	16	US-10-114-104-108
8	543.4	85.6	21646	13	US-10-087-192-910
9	533.8	84.1	2946	13	US-10-114-893-134
10	533.8	84.1	2946	15	US-10-016-249-3
11	532.8	83.9	326014	9	US-09-731-231A-3

c	12	532.8	83.9	326014	18	US-10-751-985-3	Sequence 3, Appli
	13	525.8	82.8	1136	17	US-10-632-793-25	Sequence 25, Appli
	14	522.8	82.3	283351	18	US-10-719-993-7065	Sequence 7065, Ap
	15	522.8	82.3	283351	19	US-10-741-600-17995	Sequence 17995, A
	16	522.8	82.3	1980090	18	US-10-719-993-6815	Sequence 6815, Ap
	17	522.8	82.3	1980090	19	US-10-741-600-17676	Sequence 17676, A
	18	522.6	82.3	2782	17	US-10-632-793-26	Sequence 26, Appli
	19	521	82.0	2782	17	US-10-133-036-1	Sequence 1, Appli
	20	517.4	81.5	2930	9	US-09-802-535-1	Sequence 1, Appli
	21	516.2	81.3	56093	10	US-09-873-367C-81	Sequence 81, Appli
	22	511.4	80.5	7582	17	US-10-632-793-30	Sequence 30, Appli
	23	510.6	80.4	161334	13	US-10-087-192-730	Sequence 730, App
c	24	500.8	78.9	285020	13	US-10-087-192-1666	Sequence 1666, Ap
	25	499.6	78.7	1894	9	US-09-864-761-4444	Sequence 4444, Ap
	26	495.6	78.0	3372	17	US-10-632-793-28	Sequence 28, Appli
	27	493.2	77.7	2052	17	US-10-276-774-678	Sequence 678, App
	28	482.6	76.0	22436	11	US-09-997-723-148	Sequence 148, App
	29	481.4	75.8	849	17	US-10-220-120-15	Sequence 15, Appli
	30	466	73.4	2349	17	US-10-104-047-1148	Sequence 1148, Ap
	31	449	70.7	1564	15	US-10-204-887-10	Sequence 10, Appli
	32	446.2	70.3	1393	17	US-10-220-120-17	Sequence 17, Appli
	33	434.4	68.4	7974	14	US-10-198-846-9936	Sequence 9936, Ap
	34	427.6	67.3	44063	18	US-10-322-281-718	Sequence 718, App
	35	424.6	66.9	604	13	US-10-027-632-289680	Sequence 289680,
c	36	424.6	66.9	604	17	US-10-027-632-289680	Sequence 289680,
	37	420.4	66.2	50353	19	US-10-741-600-17699	Sequence 17699, A
	38	415.6	65.4	1394	18	US-10-466-531-47	Sequence 47, Appli
	39	395.8	62.3	548	16	US-10-029-386-8554	Sequence 8554, Ap
c	40	388.4	61.2	23855	17	US-10-242-355-1175	Sequence 1175, Ap
c	41	383.4	60.4	657	13	US-10-027-632-224966	Sequence 224966,
c	42	383.4	60.4	657	17	US-10-027-632-224966	Sequence 224966,
c	43	373	58.7	439	9	US-09-764-881-183	Sequence 183, App
c	44	373	58.7	439	10	US-09-764-881-183	Sequence 183, App
c	45	373	58.7	439	17	US-10-242-747-183	Sequence 183, App

ALIGNMENTS

RESULT 1
US-08-979-847-102
; Sequence 102, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435

Fri Feb 25 16:26:30 2005

ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-979-847-102

Query Match 100.0%; Score 635; DB 8; Length 635;
Best Local Similarity 100.0%; Pred. No. 2.4e-187;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTGTTAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
DB 1 CCTGTATCTTTAACTCTCTGTTAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
QY 61 CAAATTTGTTCTTCAATGAGCACCAGATGGATGCGATCAATGAAGTCCACCGTGACCC 120
DB 61 CAAATTTGTTCTTCAATGAGCACCAGATGGATGCGATCAATGAAGTCCACCGTGACCC 120
QY 121 CTGGACCGGCTCTAGCCCATCTCCGATGTTAATGACATTAAGGACCCCTCCCGAG 180
DB 121 CTGGACCGGCTCTAGCCCATCTCCGATGTTAATGACATTAAGGACCCCTCCCGAG 180
QY 181 GAAATCTCAACTGCAAAACCCCTACTATGCCCCCAATTCAGCGGAGCAGTTAGACGGT 240
DB 181 GAAATCTCAACTGCAAAACCCCTACTATGCCCCCAATTCAGCGGAGCAGTTAGACGGT 240
QY 241 CATCAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 241 CATCAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
QY 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAATCCCTAAGCCTAGCTGGGAAGTGACT 360
DB 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAATCCCTAAGCCTAGCTGGGAAGTGACT 360
QY 361 GCATCCACCTCTAAACATGCGGCTTGCACTTAGCTCAACAGGAAATCCCTAAGCCTAGCTGGGAAGTGACT 420
DB 361 GCATCCACCTCTAAACATGCGGCTTGCACTTAGCTCAACAGGAAATCCCTAAGCCTAGCTGGGAAGTGACT 420
QY 421 ACTAAATGCTAATTTAGGCCAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
DB 421 ACTAAATGCTAATTTAGGCCAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
QY 481 AGAGCAGCGGAGGAGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGGCAACGG 540
DB 481 AGAGCAGCGGAGGAGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGGCAACGG 540
QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGCGGCTCTGTTTCACTCTATTTCCTCT 600
DB 541 CAACCCCTTTGGTCCCTCCCTTTGATGCGGCTCTGTTTCACTCTATTTCCTCT 600
QY 601 ATTAATCTTGCAACTGAAAAAAGAAAAAAGAAAAA 635
DB 601 ATTAATCTTGCAACTGAAAAAAGAAAAAAGAAAAA 635

RESULT 2
US-10-114-104-102
Sequence 102, Application US/10114104
Publication No. US20030198647A1
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
BESME, FREDERIC
BEDIN, FREDERIC

PARANHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUKER, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,104
FILING DATE: 03-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,847
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-10-114-104-102

Query Match 100.0%; Score 635; DB 16; Length 635;
Best Local Similarity 100.0%; Pred. No. 2.4e-187;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTGTTAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
DB 1 CCTGTATCTTTAACTCTCTGTTAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
QY 61 CAAATTTGTTCTTCAATGAGCACCAGATGGATGCGATCAATGAAGTCCACCGTGACCC 120
DB 61 CAAATTTGTTCTTCAATGAGCACCAGATGGATGCGATCAATGAAGTCCACCGTGACCC 120
QY 121 CTGGACCGGCTCTAGCCCATCTATGCCCCCAATTCAGCGGAGCAGTTAGACGGT 180
DB 121 CTGGACCGGCTCTAGCCCATCTATGCCCCCAATTCAGCGGAGCAGTTAGACGGT 180
QY 181 GAAATCTCAACTGCAAAACCCCTACTATGCCCCCAATTCAGCGGAGCAGTTAGACGGT 240
DB 181 GAAATCTCAACTGCAAAACCCCTACTATGCCCCCAATTCAGCGGAGCAGTTAGACGGT 240
QY 241 CATCAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 241 CATCAGCAACCTCCCAACAGCAGTCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
QY 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAATCCCTAAGCCTAGCTGGGAAGTGACT 360
DB 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAATCCCTAAGCCTAGCTGGGAAGTGACT 360

Db 43659 CATCAGCAACCTCACCACAGCACTTGGGTTTTCTGTGTGAGAGGGGACTGAGAGAC 43600
QY 301 AGGACTAGCTGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGTGACT 360
Db 43599 AGGACTAGCTGATTTCTTAGGCCAATTAAGAAATCCCTAAACCTAGCTGGGAAGTGACC 43540
QY 361 GCATCCACTCTAATACATGGGCTTGCACCTTAGCTTCACACCGCACCAATCAGAGAGCTC 420
Db 43539 GCATCCACTTTAAACACGGGCTTGCACCTTAGCTTCACACCAACCAATCAGAGAGCTC 43480
QY 421 ACTAAATGCTAATTAGGCMAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
Db 43479 ACTAAATGCTAATTAGGCMAAAACAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 43420
QY 481 AGAGCACAGCGGAGGACCAAGGATCGGGATATAAACCCAGCANTTCAGCCGGCAACGG 540
Db 43419 AGAGCACAGTGGGAGGACCAAGGATTCGAATATAAACCCAGCANTTCAGCCGGCAACGG 43360
QY 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 43359 CAACCGCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTCACTCTATTTCACCTCT 43300
QY 601 ATTAATCTTGCAACTGAAAAAATAAAAAA 632
Db 43299 ATTAATCTTGCAACTGAAAAAATAAAAAA 43268

RESULT 5
US-10-416-642-4
; Sequence 4, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: ARVIZU, Chandra
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0842 PCT
; CURRENT APPLICATION NUMBER: US/10/416,642
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/249,407
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7477736CB1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1995
; OTHER INFORMATION: a, t, c, g, or other
US-10-416-642-4

Query Match 86.0%; Score 546.4; DB 17; Length 2074;
Best Local Similarity 93.8%; Pred. No. 1.9e-159;
Matches 579; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
QY 1 CCTGTATCTTTAACTCCTCTGTTAAGTTTGTCTTTCAGAAATCAAACTGTAAACTA 60
Db 1459 CCTGTATCTTTAACTCCTCTGTTAAGTTTGTCTTTCAGAAATTCGAGCTGTAAAGCTA 1518
QY 61 CAAATTTGTTCTTCAAAATGGAGCACCAAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 1519 CAAATGTTCTTCAAAATGGAGCCCAAGATGCAGTCCATGACTTAAATCTTACCACGACCC 1578
QY 121 CTGGACGGGCTCTAGCCCATGCTCCGATGTTAATGACATTCGAAGGACCCCTCCCGAG 180
Db 1579 CTGGACGGGCTCTAGCCCATGCTCCGATGTTAATGACATTCGAAGGACCTCTCCCAAG 1638
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGATTAGAGCGGT 240

Db 1639 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGAAGCAGTTAGAGCGGT 1698
QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGGGACTGAGAGAC 300
Db 1699 CGTCAGTCAACCTCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGGGACTGAGAGAC 1758
QY 301 AGGACTAGCTGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGTGACT 360
Db 1759 AGGACTAGCTGATTTCTTAGGCCGATTAAGAAATCCCTAAGCCTAGCTGGGAAGTGACC 1818
QY 361 GCATCCACTCTAATACATGGGCTTGCACCTTAGCTTCACACCGCACCAATCAGAGAGCTC 420
Db 1819 CGTCCCACTTTAAACACGGGCTTGCACCTTAGCTTCACACCAACCAATCAGAGAGCTC 1878
QY 421 ACTAAATGCTAATTAGGCMAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
Db 1879 ACTAAATGCTAATTAGGCMAAAACAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 1938
QY 481 AGAGCACAGCGGAGGACCAAGGATCGGGATATAAACCCAGCANTTCAGCCGGCAACGG 540
Db 1939 AGAGCACAGTGGGAGGACCAAGGATTCGAATATAAACCCAGCANTTCAGCCGGCAACGG 1997
QY 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 1998 CAACCGCTTTGGGTCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 2057
QY 601 ATTAATCTTGCAACTG 617
Db 2058 ATTAATCTTGCAACTG 2074

RESULT 6
US-08-979-847-108
; Sequence 108, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIEF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787

Db 19931 GCATCCACCTTTAAACATAGGGGCTTCCACCTAGCTCACACCAACCAATCAGAGAGCTC 19990
 Qy 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG 480
 Db 19991 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG 20050
 Qy 481 AGAGCACAGGGGAGGAGCAAGGATCGGATATATAACCCAGGCAATTCGAGCCGCAACGG 540
 Db 20051 AGAGCACAGGAGGAGCAAGGATCGGATATATAACCCGCGCATTAAGCCGCAATGG 20110
 Qy 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGGCTCTGTTTCACTCTATT 593
 Db 20111 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGAGCTCTGTTTCACTCTATT 20163

RESULT 9
 US-10-114-893-134
 ; Sequence 134, Application US/10114893
 ; Publication No. US20020193567A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallie, Edward R. A.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Bowman, Michael R.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Carlin-Duckett, McKeough
 ; APPLICANT: Kelleher, Kerry S.
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: GI 6000-10A
 ; CURRENT APPLICATION NUMBER: US/10/114,893
 ; CURRENT FILING DATE: 2002-04-02
 ; EARLIER APPLICATION NUMBER: 09/413,232
 ; EARLIER FILING DATE: 1999-10-06
 ; NUMBER OF SEQ ID NOS: 321
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 134
 ; LENGTH: 2946
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-114-893-134

Query Match 84.1%; Score 533.8; DB 13; Length 2946;
 Best Local Similarity 91.5%; Pred. No. 1.9e-155;
 Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

Qy 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60
 Db 2323 CCCTGTATCTTTAACTCCTCTGTTAACTTGTCTTCCAGAACTCAAACTGTAAACTA 2382
 Qy 61 CAATTTGTTCTTCAATGAGGACACAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 120
 Db 2383 -----CAATGAGGCCCAAGATGAGTCCAGACTTAAGATCTACCGCAGACCC 2430
 Qy 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 180
 Db 2431 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 2490
 Qy 181 GAAATCTCAATGCAACACCCCTACTATGCCCCCAATTTCAGCGGGAAGCAGTTAGAGCGGT 240
 Db 2491 GAAATCTCAATGCAACACCCCTACTATGCCCCCAATTTCAGCGGGAAGCAGTTAGAGCGGT 2550
 Qy 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGAAGTGTAGAGAC 300
 Db 2551 CGTCCGCCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGATGGGGGAAGTGTAGAGAC 2610
 Qy 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCTAGCTGGGAAGGTGACT 360
 Db 2611 AGGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCTAGCTGGGAAGGTGACT 2670

Db 677 ACTAAATGCTAATCAGGCAAAACAGAGGTAAAGCAATAGCCAAATCATCTATTGCTG 736
 Qy 481 AGAGCACAGGGGAGGAGCAAGGATCGGATATATAACCCAGGCAATTCGAGCCGCAACGG 540
 Db 737 AGAGCACAGGGGAGGAGCAAGGATCGGATATATAACCCAGGCAATTCGAGCCGCAACAG 796
 Qy 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGGCTCTGTTTCACTCTATTGCTCT 600
 Db 797 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGAGCTCTGTTTCACTCTATTGCTCT 856
 Qy 601 ATTAAATCTTGCAACTGAA 619
 Db 857 ATTAAATCATGCAACTGCA 875

RESULT 8
 US-10-087-192-910
 ; Sequence 910, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 910
 ; LENGTH: 21646
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(21646)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-10-087-192-910

Query Match 85.6%; Score 543.4; DB 13; Length 21646;
 Best Local Similarity 94.8%; Pred. No. 5e-156;
 Matches 562; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60
 Db 19571 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTTCCAGAACTGTAAACTA 19630
 Qy 61 CAATTTGTTCTTCAATGAGGACACAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 120
 Db 19631 CAATTTGTTCTTCAATGAGGACCCAGATGAGTCCATGACTTAAGATCCACCGACCC 19690
 Qy 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 180
 Db 19691 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG 19750
 Qy 181 GAAATCTCAATGCAACACCCCTACTATGCCCCCAATTTCAGCGGGAAGCAGTTAGAGCGGT 240
 Db 19751 GAAATCTCAATGCAACACCCCTACTATGCCCCCAATTTCAGCGGGAAGCAGTTAGAGCGGT 19810
 Qy 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGAAGTGTAGAGAC 300
 Db 19811 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGAAGTGTAGAGAC 19870
 Qy 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCTAGCTGGGAAGGTGACT 360
 Db 19871 AAGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCTAGCTGGGAAGGTGACT 19930
 Qy 361 GCATCCACCTTAAACATGGGGCTTGAACCTTAGCTCACACCCGACCAATCAGAGAGCTC 420

```
QY 361 GCATCCACCTCTAAACATATGGGCTTGCAACTTAGCTCACACCGACCAATCAGAGAGCTC 420
Db 2671 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 2730

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCAATCTATTGCCTG 480
Db 2731 ACTAAATGCTAATTAGGCAAAACAGAGAGGTAAAGAAATAGCCAAATCAATCTATTGCCTG 2790

QY 481 AGAGCAGCGGGAGGAGGATCGGGATATAAACCCAGGCAATTCGAGCCGGCAACGG 540
Db 2791 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGTCTTCGAGCCGGCAACGG 2850

QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 2851 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 2910

QY 601 ATTAAATCTTGCAACTGAAAAAATAAAAAAAAAAAAAA 635
Db 2911 ATTAAATCTTGCAACTGAAAAAATAAAAAAAAAAAAAA 2945

RESULT 10
US-10-016-249-3
; Sequence 3, Application US/10016249
; Publication No. US20030100053A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B A1172A
; CURRENT APPLICATION NUMBER: US/10/016,249
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US/09/175,928
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-249-3
```

```
Query Match 84.1%; Score 533.8; DB 15; Length 2946;
Best Local Similarity 91.5%; Pred. No. 1.9e-155;
Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAAACTA 60
Db 2323 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAAACTA 2382

QY 61 CAAATTGTTCTTCAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 2383 -----CAATGGAGCCCAAGATGAGTCCAAAGACTAAGATCTACCGGAGACCC 2430

QY 121 CTGACCGGCTGTAGCCCATGCTCGATGTTTAAATGACATTTGAAGGACACCCCTCCCGAG 180
Db 2431 CTGACCGGCTGTAGCCCATGCTCGATGTTTAAATGACATTTGAAGGACACCCCTCCCGAG 2490

QY 181 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 2491 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 2550

QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
Db 2551 CGTCGGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 2610
```

```
QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCTAGCTGGAAAGGTGACT 360
Db 2611 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGCTAGCTGGAAAGGTGACC 2670

QY 361 GATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCGGACCAATCAGAGAGCTC 420
Db 2671 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 2730

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCAATCTATTGCCTG 480
Db 2731 ACTAAATGCTAATTAGGCAAAACAGAGAGGTAAAGAAATAGCCAAATCAATCTATTGCCTG 2790

QY 481 AGAGCAGCAGCGGAGGACAAAGATCGGGATATAAACCCAGGCAATTCGAGCCGGCAACGG 540
Db 2791 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGTCTTCGAGCCGGCAACGG 2850

QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 2851 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 2910

QY 601 ATTAAATCTTGCAACTGAAAAAATAAAAAAAAAAAAAA 635
Db 2911 ATTAAATCTTGCAACTGAAAAAATAAAAAAAAAAAAAA 2945

RESULT 11
US-09-731-231A-3/c
; Sequence 3, Application US/09731231A
; Patent No. US20020082189A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001007
; CURRENT APPLICATION NUMBER: US/09/731,231A
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 326014
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(326014)
; OTHER INFORMATION: n = A,T,C or G
US-09-731-231A-3
```

```
Query Match 83.9%; Score 532.8; DB 9; Length 326014;
Best Local Similarity 90.5%; Pred. No. 3.6e-154;
Matches 583; Conservative 0; Mismatches 52; Indels 9; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAAACTA 60
Db 170822 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAAATTTGAAGCTCTAAAACTA 170763

QY 61 CAAATTGTTCTTCAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 170762 CAAATTGTTCTTCAATGGAGCCCAAGCAGCATGATGACTAAGATCTACCGGAGACCC 170703

QY 121 CTGACCGGCTGTAGCCCATGCTCGATGTTTAAATGACATTTGAAGGACACCCCTCCCGAG 180
Db 170702 CTGACCGGCTGTAGCCCATGCTCGATGTTTAAATGACATCGNAGGACACCCCTCCAGAG 170643

QY 181 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 170642 GAAATCTCAATGACATAACCCCACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCAGT 170583

QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
Db 170582 CATCGGCCAATCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 170523
```

QY	301	AGGACTAGCTGGATTTCTTAGGCCAAACGAAGAAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACT	360
Db	170522	AGGACTAGCTGGATTTCTTAGGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	170463
QY	361	GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTACACCCGACCAATC-----	411
Db	170462	GCTTCTACCTTTAAACCCGGGCTTGCACTTAGCTACACCTAGCAATCAGGTAGGAA	170403
QY	412	AGAGAGCTCACTAAATGCTTAATTAGCCAAATAAGGAGTAAAGAAATAGCCAATCATC	471
Db	170402	AGAGAGCTCACTAAATGCTTAATTAGCCAAATAAGGAGTAAAGAAATAGCCAATCATC	170343
QY	472	TATTCCTCTGAGACACAGCGGAGGACAAGGATCGGATATATAACCCAGGCAATTCGAGC	531
Db	170342	TATCGCTGAGAACACAGTGGGAGGACATGATGGGTTATAAACCCAGGCAATTCGAGC	170283
QY	532	CGGCAACGGCAACCCCTTTGGGTCCCTCTTTGATGGGCGCTCTGTTTCACTCTA	591
Db	170282	CAGCAATGGCTACCATTTTGGGTCCCTCTTTGATGGGAGCTCTGTTTCACTCTA	170223
QY	592	TTTCACCTCTATTAATCTTGCAACTGAAAAAATAAAAAA	635
Db	170222	TTTCACCTCTATTAATCTTGCAACAGCAAAAAAATAAAAAA	170179

RESULT 13
US-10-632-793-25
; Sequence 25, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-25

Query Match 82.8%; Score 525.8; DB 17; Length 1136;
Best Local Similarity 90.7%; Pred. No. 3.9e-153;
Matches 576; Conservative 0; Mismatches 47; Indels 12; Gaps 1;

QY	1	CCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCAGAAATCAAAACTGTAAACTA	60
Db	510	CCCTGTATCTTTAACTCTCTTAACTCTTCTCTTCAGAAATCGAAGCTGTAAACTA	569
QY	61	CAAAATTTGTTCTTCAATGGAGCACAGATGGAGTCCATGATCAAGATCCACCGTGACCC	120
Db	570	-----CAAAATGGAGCCCAAGATGCAGTCCAGACTAAGATCTACCGGAGACCC	617
QY	121	CTGACCGGCTCTGCTAGCCATGCTCCGATGTTAATGACATTGAAGGACCCCTCCCGAG	180
Db	618	CTGACCGGCTCTGCTAGCCATGCTCCGATGTTAATGACATTGAAGGACCCCTCTGAG	677
QY	181	GAAATCTCAACTGCACAAACCCCTACTATATGCCCAATTCAGCGGAGAGCTAGACGGT	240
Db	678	GAAATCTCACTGTCACAACTCTACTACGCCCAATTCAGCGGAGAGCTAGACGGT	737

QY	301	AGGACTAGCTGGATTTCTTAGGCCAAACGAAGAAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACT	360
Db	170522	AGGACTAGCTGGATTTCTTAGGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	170463
QY	361	GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTACACCCGACCAATC-----	411
Db	170462	GCTTCTACCTTTAAACCCGGGCTTGCACTTAGCTACACCTAGCAATCAGGTAGGAA	170403
QY	412	AGAGAGCTCACTAAATGCTTAATTAGCCAAATAAGGAGTAAAGAAATAGCCAATCATC	471
Db	170402	AGAGAGCTCACTAAATGCTTAATTAGCCAAATAAGGAGTAAAGAAATAGCCAATCATC	170343
QY	472	TATTCCTCTGAGACACAGCGGAGGACAAGGATCGGATATATAACCCAGGCAATTCGAGC	531
Db	170342	TATCGCTGAGAACACAGTGGGAGGACATGATGGGTTATAAACCCAGGCAATTCGAGC	170283
QY	532	CGGCAACGGCAACCCCTTTGGGTCCCTCTTTGATGGGCGCTCTGTTTCACTCTA	591
Db	170282	CAGCAATGGCTACCATTTTGGGTCCCTCTTTGATGGGAGCTCTGTTTCACTCTA	170223
QY	592	TTTCACCTCTATTAATCTTGCAACTGAAAAAATAAAAAA	635
Db	170222	TTTCACCTCTATTAATCTTGCAACAGCAAAAAAATAAAAAA	170179

RESULT 12
US-10-751-985-3/C
; Sequence 3, Application US/10751985
; Publication No. US20040126861A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001007CON
; CURRENT APPLICATION NUMBER: US/10/751,985
; CURRENT FILING DATE: 2004-01-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 326014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(326014)
; OTHER INFORMATION: n = A, T, C or G
US-10-751-985-3

Query Match 83.9%; Score 532.8; DB 18; Length 326014;
Best Local Similarity 90.5%; Pred. No. 3.6e-154;
Matches 583; Conservative 0; Mismatches 52; Indels 9; Gaps 1;

QY	1	CCCTGTATCTTTAACTCTGTTAAGTTTGTCTCTTCAGAAATCAAAACTGTAAACTA	60
Db	170822	CCCTGTATCTTTAACTCTGTTAAGTTTGTCTCTTCAGAAATGAAGCTCTAAACTA	170763
QY	61	CAAAATTTGTTCTTCAATGGAGCACAGATGGAGTCCATGATCAAGATCCACCGTGACCC	120
Db	170762	CAAAATTTGTTCTTCAATGGAGCCCGACAGCAGTCCATGATCAAGATCTACCGAGACCC	170703
QY	121	CTGACCGGCTCTGCTAGCCATGCTCCGATGTTAATGACATTGAAGCACCCTCCCGAG	180
Db	170702	CTGACCGGCTCTGCTAGCCATGCTCCGATGTTAATGACATCGAAGCAACCCCTCCAGAG	170643
QY	181	GAAATCTCACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT	240
Db	170642	GAAATCTCACTGCATACCCCACTACGCCCAATTCAGCGGGAAGCAGTTAGACAGT	170583
QY	241	CATCAGGCCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGAC	300
Db	170582	CATCGGCCATCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGTACTCAGAGAC	170523

Qy	241	CATCAGCGCAACCTCCCCAAACAGCACTCTGGGTTTTCTGTTGAGAGGGGGAGACTGAGAGAC	300
Db	738	CGTCGGGCAACCTCCCCAAACAGCACTCTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC	797
Qy	301	AGGACTAGCTGGATTTCCCTAGGCGCAACGAGAATCCCTAAGCCTAGCTGGGAAGGTCACT	360
Db	798	AGGACTAGCTGGATTTCTAGGCTGACTAAGATCCCTAAGCCTAGCTGGGAAGGTGACC	857
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420
Db	858	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGTTACACCTGACCAATCAGAGAGCTC	917
Qy	421	ACTAAATGCTAATATTAGGCAAAAAATAGGAGGTAAAGAAATAGCCAACTATTTGCTGT	480
Db	918	ACTAAAAATGCTAATATTAGGCAAAGACAGGAGGTAAAGAAATAGCCAAATCATCTATTGGCATG	977
Qy	481	AGAGCACAGCGGGAGGAGCAAGGATCGGGATATAAACCCAGGCATTCGAGCGGCAACGG	540
Db	978	AGAGCACAGCGAGGAGGAGCAATGATCGGGATATAAACCCAGCTCTTCGAGCGGCAACGG	1037
Qy	541	CAACCCCTCTTTGGGTCCCTCCCTTTGTATGGGCGCTGTGTTTTCACTCTATTTCACTCT	600
Db	1038	CAACCCCTCTTTGGGTCCCTCCCTTTGTATGGGAGCTGTGTTTTCACTCTATTTCACTCT	1097
Qy	601	ATTTAAATCTTCGAACCTGAAAAAATAAAAAA	635
Db	1098	ATTTAAATCTTCGAGCTGCGAAAAAATAAAAAA	1132

```

RESULT 14
US-10-719-993-7065
; Sequence 7065, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARCILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7065
; LENGTH: 283351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(283351)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tab
US-10-719-993-7065

```

Db	175045	TGGGTTTTCTGTTGAGAGGGGACCTGAGAGACAGGACTAGCTGGAATTCCTAGGCCGA	175104
Qy	327	CGAAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACTGATCCACCTCTAAACATGGGGCTTGG	386
Db	175105	CTAAGAATCCCTAAGCCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG	175164
Qy	387	CAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAATGCTAAATTAGGCACAAAATA	446
Db	175165	CAACTTAGCTCACAC-----CCAACAGAGAGTTCCTAAATGCTAAATTAGGCACAAAATA	175220
Qy	447	GGAGTAAAGAAATAGCCAAATCATCTATTGGCTGAGAGCACAGCGGGAGGGAACAAGGATC	506
Db	175221	GGAGTACAGAAATAGCCAAATCATCTATTGGCTGAGAGCACAGCGGGAGGGAACAAGGATC	175280
Qy	507	GGGATATAACCCAGGCATTTCAGCGGCACAGGCACCCCTTTGGGTCCCTCCCTTT	566
Db	175281	AGGATATAAACCCAGGCATTTCAGCTGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT	175340
Qy	567	GTATGGGGCTCTGTTTTCACCTCTATTATTCACCTCTATTAAATCTTGCACACTGAAAAAAA	626
Db	175341	GTATGGGAGCTCTGTTTCA--CTGTTTCACCTCTATTAAATCTTGCACACTGCAAAAAGA	175398
Qy	627	AAAAAAA 634	
Db	175399	AAAAAAA 175406	

```

RESULT 15
US-10-741-600-17995
; Sequence 17995, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: fastSEQ for Windows Version 4.0
; SEQ ID NO 17995
; LENGTH: 283351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(283351)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Table
US-10-741-600-17995

```

	Query Match	82.3%;	Score 522.8;	DB 18;	Length 283351;
	Best Local Similarity	93.4%;	Pred. No. 4.4e-151;		
	Matches 568;	Conservative 1;	Mismatches 33;	Indels 6;	Gaps 2;
Qy	27	GTTTGTCTCTTCGAGATCAAAAATCTGTAARAATCAAAATGTTCTTCAAAATGGAGCACCA	86		
Db	174805	GTTTGTCTCTTCCAGAAATTGAAATCTGTAARAATCAAAATGGTTCTTCAAAATGGAGGCCCA	174864		
Qy	87	GATGGAGTCCATGACTAAGATCTCACCGTGGACCCCTGGACCGGCGCTGTGTAGCCCATGCTC	146		
Db	174865	GATGCAGTCCATGACTAAGATCTCACCGCAGACCCCTGGACYGCGCTGCTAGCCCATGCTC	174924		
Qy	147	CGATGTTAATGACATTTGAAGGCAACCCCTCCGAGGAAATCTCAACTGGACACCCCTACT	206		
Db	174925	CAATGTTAATGACATTTGAAGGCAACCCCTTCGAGGAGAAATCTCAACTACACAACACCTACT	174984		
Qy	207	ATGCCCCAAATTCAGGGGAAGCAGTTAGAGGGGTTCATCAGGCCAACCTCCCCAACAGCACT	266		
Db	174985	ATGCCCCAAATTCAGGAGGAGCTGTTAGAGCAGTCATTGGGCCAACCTCCCCAACAGCACT	175044		
Qy	267	TGGGTTTTTCTGTTGAGAGGGGGGCACTGAGAGACGAGCACTAGCTGGATTTCTCTTAGGCCAA	326		

Query Match	82.3%	Score	522.8	DB	19	Length	283351
Best Local Similarity	93.4%	Pred. No.	4.4e-151				
Matches	568	Conservative	1	Mismatches	33	Indels	6
Qy	27	GTTTGTCTCTTCCAGAACTCAAAA	CTGTAAAACTACAAATGTGTTCTTCAAATGGAGACCA	86			
Db	174805	GTTTGTCTCTTCCAGAAATTGAA	CTGTAAAACTACAAATGTGTTCTTCAAATGGAGCCCA	174864			
Qy	87	GATGGAGTCCATGACTTAAGATCC	ACCGTGGACCGGCTGTAGCCCATGCTC	146			
Db	174865	GATGCAGTCCATGACTTAAGATCC	ACCGCAGACCCCTGACATGGCTGTAGCCCATGCTC	174924			
Qy	147	CGATGTTTAATGACATTTGAAGC	ACCCCTCCGAGGAAATCTCAACTGCACAAACCCTACT	206			
Db	174925	CAATGTTTAATGACATGGAAAGCA	CCCTCCGAGGAAATCTCAACTACACACACTACT	174984			
Qy	207	ATGCCCCAAATTCAGCGGGAAGC	ATTAGAGCGGTATCAGGCCAACCTCCCCAACAGCACT	266			
Db	174985	ATGCCCCAAATTCAGCAGGAAGCT	TTTAGAGCAGTCATTGGCCAACTCCCCAACAGCACT	175044			
Qy	267	TGGGTTTTCTTTGAGAGGGGGG	AGCTCAGAGACAGGACTAGCTGGATTTCTTAGGCCAA	326			

Db 175045 TGGGTTTCTGTTGAGAGGGGACTGAGACAGGACTAGCTGGATTTCTTAGGCCGA 175104

Qy 327 CGAAGAATCCCTAAGCCTAGCTGGAAAGGTGACTGCATCCACCTCTAAACATGGGGCTTG 386

Db 175105 CTAAGAATCCCTAAGCCTAGCTGGAAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 175164

Qy 387 CAACTTAGCTCACACCGGACCAATCAGAGAGCTCACTAAATGCTAAATTAGGCAAAATA 446

Db 175165 CAACTTAGCTCACAC-----CCAACAGAGAGTTTCACTAAATGCTAAATTAGGCAAAATA 175220

Qy 447 GGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGCACAGCGGGAGGACAGGATC 506

Db 175221 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGAGAGCACAGCGGGAGGACAGGATC 175280

Qy 507 GGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 566

Db 175281 AGGATATAAACCCAGGCAATTCGAGCTGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 175340

Qy 567 GTATGGGCGCTGTGTTTCACTCTATTTCACCTCTATTAAATCTTGCAACTGAAAAAAA 626

Db 175341 GTATGGGAGCTGTGTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCAAAAAGA 175398

Qy 627 AAAAAAA 634

Db 175399 AAAAAAA 175406

RESULT 16

US-10-719-993-6815

; Sequence 6815, Application US/10719993

; Publication No. US20040265849A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; FILE REFERENCE: CL001496

; CURRENT APPLICATION NUMBER: US/10/719,993

; CURRENT FILING DATE: 2003-11-24

; NUMBER OF SEQ ID NOS: 55342

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6815

; LENGTH: 1980090

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(1980090)

; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-10)

Query Match 82.3%; Score 522.8; DB 18; Length 1980090;

Best Local Similarity 93.4%; Pred. No. 1.1e-150;

Matches 568; Conservative 1; Mismatches 33; Indels 6; Gaps 2;

Qy 27 GTTTGTCTCTCCAGAAATCAAACTGTAAACTACAAATGTTCTTCAAATGGAGCACCA 86

Db 1777083 GTTTGTCTCTCCAGAAATGAAACTGTAAACTACAAATGTTCTTCAAATGGAGCCCCA 1777142

Qy 87 GATGGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGCTGTAGCCCATGCTC 146

Db 1777143 GATGGAGTCCATGACTAAGATCCACCGAGACCCCTGGACCGGCTGTAGCCCATGCTC 1777202

Qy 147 CGATGTTAATGACATTTGAAGGACCCCTCCGAGGAAATCTCAACTGCACACCCCTACT 206

Db 1777203 CAATGTTAATGACATTTGAAGGACCCCTCCGAGGAAATCTCAACTGCACACCCCTACT 1777262

Qy 207 ATGCCCAATTCAGCGGGAAGAGTGTAGAGCGGTCACTCAGCCCACTCCCCCAACAGCACT 266

Db 1777263 ATGCCCAATTCAGCGGGAAGAGTGTAGAGCGGTCACTCAGCCCACTCCCCCAACAGCACT 1777322

Qy 267 TGGGTTTTCTGTTAGAGGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAA 326

Db 1777323 TGGGTTTTCTGTTAGAGGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGA 1777382

Qy 327 CGAAGAATCCCTAAGCCTAGCTGGAAAGGTGACTGCATCCACCTCTAAACATGGGGCTTG 386

Db 1777383 CTAAGAATCCCTAAGCCTAGCTGGAAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 1777442

Qy 387 CAACTTAGCTCACACCGGACCAATCAGAGAGCTCACTAAATGCTAAATTAGGCAAAATA 446

Db 1777443 CAACTTAGCTCACAC-----CCAACAGAGAGTTTCACTAAATGCTAAATTAGGCAAAATA 1777498

Qy 447 GGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGCACAGCGGGAGGACAGGATC 506

Db 1777499 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGAGAGCACAGCGGGAGGACAGGATC 1777558

Qy 507 GGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 566

Db 1777559 AGGATATAAACCCAGGCAATTCGAGCTGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 1777618

Qy 567 GTATGGGCGCTGTGTTTCACTCTATTTCACCTCTATTAAATCTTGCAACTGAAAAAAA 626

Db 1777619 GTATGGGAGCTGTGTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCAAAAAGA 1777676

Qy 627 AAAAAAA 634

Db 1777677 AAAAAAA 1777684

RESULT 17

US-10-741-600-17676

; Sequence 17676, Application US/10741600

; Publication No. US20050026169A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; FILE REFERENCE: CL001499

; CURRENT APPLICATION NUMBER: US/10/741,600

; CURRENT FILING DATE: 2003-12-22

; NUMBER OF SEQ ID NOS: 73997

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17676

; LENGTH: 1980090

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(1980090)

; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-10)

Query Match 82.3%; Score 522.8; DB 19; Length 1980090;

Best Local Similarity 93.4%; Pred. No. 1.1e-150;

Matches 568; Conservative 1; Mismatches 33; Indels 6; Gaps 2;

Qy 27 GTTTGTCTCTCCAGAAATCAAACTGTAAACTACAAATGTTCTTCAAATGGAGCACCA 86

Db 1777083 GTTTGTCTCTCCAGAAATGAAACTGTAAACTACAAATGTTCTTCAAATGGAGCCCCA 1777142

Qy 87 GATGGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGCTGTAGCCCATGCTC 146

Db 1777143 GATGGAGTCCATGACTAAGATCCACCGAGACCCCTGGACCGGCTGTAGCCCATGCTC 1777202

Qy 147 CGATGTTAATGACATTTGAAGGACCCCTCCGAGGAAATCTCAACTGCACACCCCTACT 206

Db 1777203 CAATGTTAATGACATTTGAAGGACCCCTCCGAGGAAATCTCAACTGCACACCCCTACT 1777262

Qy 207 ATGCCCAATTCAGCGGGAAGAGTGTAGAGCGGTCACTCAGCCCACTCCCCCAACAGCACT 266

Db 1777263 ATGCCCAATTCAGCGGGAAGAGTGTAGAGCGGTCACTCAGCCCACTCCCCCAACAGCACT 1777322

Qy 267 TGGGTTTTCTGTTAGAGGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAA 326

Db 1777323 TGGGTTTTCTGTTAGAGGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGA 1777382

QY 327 CGAAGATCCCTAGCCTAGCTGGAGGTGACTGATCCACCTCTAAACATGGGGCTTG 386
Db 1777383 CTAAGATCCCTAGCCTAGCTGGAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 1777442
QY 387 CAATCTAGCTCACACCGGACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAATA 446
Db 1777443 CAATCTAGCTCACAC-----CCAACGAGAGGTTCACCTAAAATGCTAATAGGCAAAATA 1777498
QY 447 GGAGTTAAAGAAATAGCAATCATCTATTGCTTGGAGGACACGGGGAGGACAGAGATC 506
Db 1777499 GGAGTTACAGAAATAGCAATCATCTATTGCTTGGAGGACACGGGGAGGACAGAGATC 1777558
QY 507 GGGATATAAACCCAGGCAATCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 566
Db 1777559 AGGATATAAACCCAGGCAATCGAGCTGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 1777618
QY 567 GTATGGCGCTCTCTTTTCACTCTATTCTATTCTATTAAATCTTTGCAACTGAAAAAAA 626
Db 1777619 GTATGGGAGCTCTGTTTCA--CTGTTTCACTCTATTAAATCTTTGCAACTGCAAAAAAGA 1777676
QY 627 AAAAAAA 634
Db 1777677 AAAAAAA 1777684

RESULT 18
US-10-632-793-26
; Sequence 26, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-26

Query Match 82.3%; Score 522.6; DB 17; Length 2782;
Best Local Similarity 90.4%; Pred. No. 5.9e-152;
Matches 574; Conservative 0; Mismatches 49; Indels 12; Gaps 1;
QY 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTGCTCTTCCAGAACTAAACTGTAATACTA 60
Db 2158 CCCTGTATCTTTGACCTCTCTTTGTTAACTCTCTTCCAGAACTGTAATACTA 2217
QY 61 CAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 2218 -----CAATGGAGCCCAAGATGAGTCCCAAGATTAAGATCTACGGAGACCC 2265
QY 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGATTAAGGACACCCCTCCCGAG 180
Db 2266 CTGGACCGGCTCTAGCCCATGCTCTATGCCCCAATTCAGCGGGAAGCAGTTAGCGGT 240
QY 181 GAAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGCGGT 2385
Db 2326 GAAATCTCAGCTGCAACACCTCTACTACGCCCAATTCAGCGGGAAGCAGTTAGCGGT 2385
QY 241 CATCAGCCAACTCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300

Db 2386 GGTGGGCCAACCTCCCAACAGCAGCACTTAGGTTTTCTGTTAGATGGGGACTGAGAGAC 2445
QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCTTAAGCCTAGCTGGGAAGGTGACT 360
Db 2446 AGGACTAGCTGGATTTCTTAGGCCAACGAAGATCCTTAAGCCTAGCTGGGAAGGTGACC 2505
QY 361 GCATCCACCTCTAAACATGGGGCTTCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 2506 ACATCCACCTTTAAACACACGGGGCTTCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 2565
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCGCTG 480
Db 2566 ACTAAATGCTAATTAGGCAAAACAGAGAGGTAAAGAAATAGCAATCATCTATTGCGCTG 2625
QY 481 AGAGCAGAGCGGAGGAGCAAGGATCGGATATAAAACCCAGGCAATTCGAGCCCGCAACGG 540
Db 2626 AGAGCAGAGCAGAGGAGGACATGATCGGATATAAAACCCAGGTTTTCGAGCCCGCAACGG 2685
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 2686 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 2745
QY 601 ATTAATCTTTGCAACTGCAAAAAA 635
Db 2746 ATTAATCTTTGCAACTGCAAAAAA 2780

RESULT 19
US-10-133-036-1
; Sequence 1, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
; APPLICANT: Conrad, Bernard
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Human endogenous retrovirus
US-10-133-036-1

Query Match 82.0%; Score 521; DB 17; Length 2782;
Best Local Similarity 90.2%; Pred. No. 1.9e-151;
Matches 573; Conservative 0; Mismatches 50; Indels 12; Gaps 1;
QY 1 CCCTGTATCTTTAACTCTCTTTGTTAAAGTTGCTCTTCCAGAACTAAACTGTAATACTA 60
Db 2158 CCCTGTATCTTTGACCTCTCTTTGTTAACTCTCTTCCAGAACTGTAATACTA 2217
QY 61 CAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 2218 -----CAATGGAGCCCAAGATGAGTCCCAAGATTAAGATCTACGGAGACCC 2265
QY 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGATTAAGGACACCCCTCCCGAG 180
Db 2266 CTGGACCGGCTCTAGCCCATGCTCTATGCCCCAATTCAGCGGGAAGCAGTTAGCGGT 240
QY 181 GAAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGCGGT 2385
Db 2326 GAAATCTCAGCTGCAACACCTCTACTACGCCCAATTCAGCGGGAAGCAGTTAGCGGT 2385
QY 241 CATCAGCCAACTCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
Db 2386 GGTGGGCCAACCTCCCAACAGCAGCACTTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 2445

301	AGGACTAGCTGGATTTCCTAGGCCAACGAAATCCCTTAAGCTTAGCTGGGAGGTC	360
2446	AGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCTTAAGCCCTAGGTGGGAAGTGACC	2505
361	GCATCCACCTCTAAACATGCGGCTTGCACCTCAGCCGACCAATCAGAGAGCTC	420
2506	ACATCCACCTTTAAACACGGGGCTTGCACCTTAGCTCACCTGACCAATCAGAGAGCTC	2565
421	ACTAAATGCTAATTAGGCCAAAAATAGGAGGTAAAGAAATAGCCAATCATTTATGCGCTG	480
2566	ACTATAATGCTAATTAGGCCAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATGCGCTG	2625
481	AGAGCACAGCGGAGGGACAGGATCGGATATAAACCAGGCAATCGAGCCGCGCAACGG	540
2626	AGAGCACAGCAGGAGGGACAATGATCGGATATAAACCCAAGTTTTCAGCCGCGCAACGG	2685
541	CAACCCCTTTGGGTCCCTTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACTCT	600
2686	CAACCCCTTTGGGTCCCTTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACTCT	2745
601	ATTAAATCTTGCACCTGAAGAAAAA	635
2746	ATTAAATCTTGCACCTGCAAAAAA	2780

```

RESULT 20
US-09-902-535-1
; Sequence 1, Application US/0902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: Mi, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preeclampsia and gestational trophoblast
; TITLE OF INVENTION: disorders
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930) ... (2546)
US-09-902-535-1

```

Query Match	81.5%	Score 517.4	DB 9	Length 2930
Best Local Similarity	91.4%	Pred. No. 2.5e-150		
Matches 564	Conservative	0	Mismatches 41	Indels 12
Gaps				
Qy	1	CCCTGTATCTTTAAACCTCTTGGTTTGTCTCTTTCCAGAAATCAAAAACGTGTAAAACTA	60	
Db	2325	CCCTGTATCTTTAAACCTCTTGGTTTAACTTTGTCTCTTTCCAGAATCGAAGCTGTAAAACTA	2384	
Qy	61	CAAAATGTTCTTTCAAATGGAGACACAGATGGAGTCCATGACTAGATCCACCGTGGACCC	120	
Db	2385	-----CAAATGGAGCCCAAGATGCAGTCCAAGACTAGATCTCCCGCAGACCC	2432	
Qy	121	CTGGACCGGCGCTGTAGCCCAATGCTCCGAGTGTTAATGACATTTGAAGGCACCCCTCCCGAG	180	
Db	2433	CTGGACCGGCGCTGTAGCCCAACGATCTGATGTTAATGACATCAAAAGGCACCCCTCCTGAG	2492	
Qy	181	GAAAATCTCAATGCACACACCCCTACTATGCCCCCAATTACAGCGGGAAGCAGTGTAGAGCGGT	240	
Db	2493	GAAAATCTCAGTGCACACACCTCTTACTACGCCCCCAATTCAGCAGGAAGCAGTGTAGAGCGGT	2552	
Qy	241	CATCAGCCCAACCTCCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAC	300	

Db	2553	CGTCGGCCAACTCCCCAAACAGCACATTAGGTTTTCTGTTGAGATGGGGACTGAGAGAC	2612
Qy	301	AGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCCTAGCTCGGAAGGTGACT	360
Db	2613	AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACC	2672
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTGCACACCCGACCAATCAGAGAGCTC	420
Db	2673	ACATCCACCTTTAAACACGGGGCTTGGCACTTAGCTCACACCTGACCAATCAGAGAGCTC	2732
Qy	421	ACTTAAATGCTAAATTTAGGCAAAAAATAGGAGGTAAAGAAATAGCCAATCATTTGCGCTG	480
Db	2733	ACTTAAATGCTAAATTTAGGCAAAAAACAGAGGTAAGAAATAGCCAATCATTTGCGCTG	2792
Qy	481	AGAGCACAGCGGGAGGGACAAGGATCGGGATATAAACCCAGGCATTTGAGCGCGGCAACGG	540
Db	2793	AGAGCACAGCAGGAGGGACATGATCGGGATATAAACCCCAAGTCTTCGAGCGCGCAACGG	2852
Qy	541	CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTTCACTCTATTATTTCACTCT	600
Db	2853	CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTATTTCACTCT	2912
Qy	601	ATTAAATCTTGCAACTG	617
Db	2913	ATTAAATCTTGCAACTG	2929

```

RESULT 21
US-09-873-367C-81
; Sequence 81, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-81

```

	Query Match	81.3%	Score 516.2	DB 10	Length 56093
	Best Local Similarity	91.1%	Pred. NO. 2.4e-149		
	Matches 564	Conservative	Mismatches 0	Indels 12	Gaps 1
QY	1	CCCTGTATCTTTAACTCCTTGTAGTTGTCTCTTCAGAAATCAAAACTCTAAAACTA	60		
DB	37274	CCCTGTATCTTTAACTCCTTGTAGTTGTCTCTTCAGAAATCGAAGCTGTAAACTA	37333		
QY	61	CAAAATTGTTCTTTCAAATGGAGACACAGATGGAGTCCACTAAAGATCCACCGTGGACCC	120		
DB	37334	-----CAAATGGAGACCAAGATGCAGTCCCAAGACTAAGATCTACCGCAGACCC	37381		
QY	121	CTGGACCGGCGCTGCTAGCCCAATGCTCGGATGTTAATGACATGAAGGCCACCCCTCCCGAG	180		


```
Db 37382 CTGGACGGCTCTAGCCACGATCTGATGTTAATGACATCAAGACACCCCTCTGAG 37441
Qy 181 GAAATCTCAACTGCACACCCCTACTATGCCCCAATTCAGCGGAAGCAGTTAGAGCGGT 240
Db 37442 GAAATCTCAGTCGACAAACCTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 37501
Qy 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 37502 CTGGGCAACCTCCCAACAGCAGCTTAGGTTTTCTGTTGAGATGGGAGCTGAGAGAC 37561
Qy 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 37562 AGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCTAAGCCTAGCTGGGAAGGTGACC 37621
Qy 361 GATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 37622 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 37681
Qy 421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
Db 37682 ACTAAATGCTAATTAGGCAAGACAGAGGTAAAGAAATAGCAATCATCTATTGCTG 37741
Qy 481 AGAGCAGCGGGAGGACGAAGATCGGGATATAAACCCAGGATTCGAGCCGGCAACGG 540
Db 37742 AGAGCAGCAGGAGGACGAATGATCGGATATAAACCCAGGCTTCGAGCCGGCAACGG 37801
Qy 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
Db 37802 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 37861
Qy 601 ATTAATCTTGCACCTGAA 619
Db 37862 ATTAATCTTGCACCTGCA 37880
```

RESULT 22

```
US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauclia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u

Query Match 80.5%; Score 511.4; DB 17; Length 7582;
Best Local Similarity 89.5%; Pred. No. 3e-148;
Matches 552; Conservative 11; Mismatches 42; Indels 12; Gaps 1;

Qy 1 CCTGTATCTTTAACCTCTCTGTTAAGTTGCTCTTCCAGAAATCAAACTGTAAACTA 60
Db 6976 CCTGTATCTTTTRACCTCTCTGTTAATCTTCTTCAGAAATCGAAGCTGTAAACTA 7035

Qy 61 CAAATTTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGATGAATCAACCGTGGACCC 120
Db 7036 -----CAAATGGAGCCAGATGCACTCCAAGACTAAGATCTACCGCAGACCC 7083

Qy 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
Db 7084 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 7143

Qy 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 240
Db 7144 GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203

Qy 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 7204 SGTCGGCAACCTCCCAACAGCAGCTTAGGTTTTCTGTTGAGATGGGAGCTGAGAGAC 7263

Qy 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 7264 AGGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCVTAAGCCTAGSTGGGAAGGTGACC 7323

Qy 361 GATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTTAGCTAGCTAGCTAGCTAGCTAGCT 7383

Qy 421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
Db 7384 ACTAAATGCTAATTAGGCAAGACAGAGGTAAAGAAATAGCAATCATCTATTGCTG 7443

Qy 481 AGAGCAGCGGGAGGACGAAGATCGGGATATAAACCCAGGATTCGAGCCGGCAACGG 540
Db 7444 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAGGATTCGAGCCGGCAACGG 7503

Qy 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
Db 7504 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 7563
```

Fri Feb 25 16:26:30 2005

```
Qy 601 ATTAATCTTGAACGTG 617
Db 7564 ATTAATCTTGCARCTG 7580

RESULT 23
US-10-087-192-730/c
; Sequence 730, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 730
; LENGTH: 161334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(161334)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-730

Query Match 80.4%; Score 510.6; DB 13; Length 161334;
Best Local Similarity 90.0%; Pred. No. 2,2e-147;
Matches 575; Conservative 0; Mismatches 49; Indels 15; Gaps 2;

Qy 1 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
Db 100611 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAAGCTA 100552

Qy 61 CAAATGTTCTTCAAAATGAGCAGACGATGGATCCATGACTTAAGATCCACCGTGACCC 120
Db 100551 CAAATGTTCTTCAAAATGAGCAGACGATGGATCCATGACTTAAGATCCACCGTGACCC 100492

Qy 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGGACCCCTCCCGAG 180
Db 100491 CTGGACCGGCTCTAGCCCATGCTCCCAATTTGTAATGATATCGAAGCCACCCCTCCCGAG 100432

Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCGCGAGGAGCAGTTAGACGGT 240
Db 100431 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCGCGAGGAGCAGTTAGACGGT 100372

Qy 241 CATCAGCCACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGGACTTGAGAGAC 300
Db 100371 CGTCAGCCACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGGACTTGAGAGAC 100312

Qy 301 AGGACTAGCTGGATTTCTAGGCGCAACGAGAAATCCCTTAAGCCTAGCTGGGAAGGTGACT 360
Db 100311 AGGACTAGCTGGATTTCTAGGCGCGACTAAGAAATCCCAAGCCTAGCTGGGAAGGTGACC 100252

Qy 361 GCATCCACCTCTAAACATGCGGCTTGCACCTTAGCTCACACCCGACCAATC 411
Db 100251 ACATCCACCTTTAAACATGCGGCTTGCACCTTAGCTCACACCCGACCAATCAGGTAGTAA 100192

Qy 412 AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATC 471
Db 100191 AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATC 100132

Qy 472 TATTGCTGAGAGCAGACCGGGAGGCAAGGATTCGGGATATAAACCCAGGCAATTCGAGC 531
Db 100131 TATCGCTGAGAGCAGACCGGGAGGCAATGATTCGGGATATAAACCCAGGCAATTCAGC 100072

Qy 532 CGCAACGCGCAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTTCACCTCTA 591
Db 100071 CGCAACGCGCTACCTCTTTTGGTCCCTCCCTTTGATGGAGCTCTCTCT-----GT 100018

Qy 592 TTTCACTCTATTAAATCTTGCAACTGAAAAAATA 630
Db 100017 CTTCACCTCTATTAAATATTGCAACTGCAAAAAAATA 99979

RESULT 24
US-10-087-192-1666
; Sequence 1666, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1666
; LENGTH: 285020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(285020)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1666

Query Match 78.9%; Score 500.8; DB 13; Length 285020;
Best Local Similarity 89.3%; Pred. No. 3.3e-144;
Matches 577; Conservative 0; Mismatches 57; Indels 12; Gaps 3;

Qy 1 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
Db 279486 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 279545

Qy 61 CAAATGTTCTTCAAAATGAGCAGACGATGGATCCATGACTTAAGATCCACCGTGACCC 120
Db 279546 CAAATGTTCTTCAAAATGAGCAGACGATGGATCCATGACTTAAGATCCACCGTGACCC 279605

Qy 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGGACCCCTCCCGAG 180
Db 279606 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTGAATGACATCAAAAGGACCCCTCTGAG 279665

Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGACGGT 240
Db 279666 GAAATCTCAACTGCACGACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGACGGT 279725

Qy 241 CATCAGCCACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGGACTTGAGAGAC 300
Db 279726 CATTCGCCAACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGTTGGGACTTGAGAGAC 279785

Qy 301 AGGACTAGCTGGATTTCTAGGCGCAACGAAAG--AATCCCTTAAGCCTAGCTGGGAAGGTGA 358
Db 279786 AGGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGA 279845

Qy 359 CTGCATCCACCTCTTAAACATGCGGCTTGCACCTTAGCTCACACCCGACCAATC----- 411
Db 279846 CGGCATCCACCTTTAAACACGCGGCTTGCACCTTAGCTCACACCCGACCAATCAGGAGT 279905

Qy 412 --AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATAGCCAAATCA 469
Db 279906 AAGAGAGCTCACTAAATGCTAAATAGGCTAAACAGAGGTTAAAGAAATAGCCAAATCA 279965
```


PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-28

Query Match 78.0%; Score 495.6; DB 17; Length 3372;
Best Local Similarity 92.2%; Pred. No. 1.7e-143;
Matches 522; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 70 CTTCAAAATGGAGCACCAGATGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGG 129
Db 2802 CTACAAATGGAGCCCAAGATGAGTCCCAAGACTAAGATCTACCGCAGACCCCTGGACCGG 2861

QY 130 CTTGCTAGCCCATGCTCCCATGTTAATGACATTTGAAGGCACCCCTCCCGAGGAAATCTCA 189
Db 2862 CTTGTTAGCCCATGCTCCCATGTTAATGACATTTGAAGGCACCCCTCCCGAGGAAATCTCA 2921

QY 190 ACTGCAAAACCCCTACTATGCCCCCAATTCAGCGGGAGCAGTTAGAGCGGTCAATCAGCCA 249
Db 2922 GCTGCAACACCTCTACTACGCCCAATTCAGCGGAGCAGTTAGAGCGGTCTGCGCCA 2981

QY 250 ACCTCCCAACAGCAGCTTGGTTTTCCTGTTGAGAGGGGGGACTGAGACAGGACTAGC 309
Db 2982 ACCTCCCAACAGCAGCTTGGTTTTCCTGTTGAGATGGGGGACTGAGACAGGACTAGC 3041

QY 310 TGGATTTCCTAGGCCCAAGAAATCCCTAAGCTAGCTGGGAAGTGTGCTCATCCACC 369
Db 3042 TGGATTTCCTAGGCTGATTAAGAAATCCCTAAGCTAGCTGGGAAGTGTGCTCATCCACC 3101

QY 370 TCTAAACATGGGCTTGGCACTTAGCTACACCGGACCAATCAGAGAGCTCACTAAATG 429
Db 3102 TTTAAACACGGGCTTGGCACTTAGCTACACCGGACCAATCAGAGAGCTCACTAAATG 3161

QY 430 CTAATTTAGGCAGAAATAGGAGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGAGCAG 489
Db 3162 CTAATTTAGGCAGAAATAGGAGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGAGCAG 3221

QY 490 CGGAGGAGCAAGGATCGGATATAAACCCAGGATTCGAGCGGCAACGGCAACCCCT 549
Db 3222 CAGGAGGAGCAATGATCGGATATAAACCCAGGATTCGAGCGGCAACGGCAACCCCT 3281

QY 550 TTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTTCACCTATTTCACCTATTAAATCT 609
Db 3282 TTGGGTCCTCCCTCTTGTATGGAGCTGTTTTCACCTATTTCACCTATTAAATCT 3341

QY 610 TGCACCTGAAATGAAAAAAAAAAAAA 635
Db 3342 TGCACCTGCAAAAAAAAAAAAAA 3367

RESULT 27
US-10-276-774-678
; Sequence 678, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 678
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-678

Query Match 77.7%; Score 493.2; DB 17; Length 2052;
Best Local Similarity 91.2%; Pred. No. 7.7e-143;
Matches 549; Conservative 0; Mismatches 43; Indels 10; Gaps 2;

QY 1 CCCTGTATCTTTAAACCTCTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAACTA 60
Db 993 CCCTGTATCTTTAAACCTCTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAAAGCTGTAAACAA 1052

QY 61 CAAATTTCTTCAAAATGGAGCACCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 1053 CACATCTCTTCAAAATGGAGCACCAGATGAGTCCATGACTAAGATCTACCGGGATCC 1112

QY 121 CTGACCGGCTGCTAGCCCATGCTCGATGTTAATGACATTTGAAGGCACCCCTCCCGAG 180
Db 1113 CTGACCGGCTGCTAGCCCATGCTCGATGTTAATGACATTTGAAGGCACCTCTCTCCGAG 1172

QY 181 GAAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
Db 1173 GAAATCTCAAAATGCAACACCCCTACTATGCCCAATTTAAGCAGGAGCAGTTGAGCAGT 1232

QY 241 CATCAGCAACCTCCCAACAGCAGCTTGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300
Db 1233 CGAGCGCAACCTCCCAACAGCAGCTTGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 1292

QY 301 AGGACTAGCTGATTTCTTAGGCCAAACGAGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 360
Db 1293 AGGACTAGCTGATTTCTTAGGCCGATAGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 1352

QY 361 GCATCCACCTCTAAACATGCGGCTTGCAATTTAGCTCTACACCCACCCATC----- 411
Db 1353 GTATCCACCTTTAAACACAGGGCTTGCAATTTAGCTCTACACCCACCCATCAGGTAGTAA 1412

QY 412 AGAGAGCTCACTAAATGCTAATTTAGGCAGAAATAGGAGGTAAAGAAATAGCCCAATCATC 471
Db 1413 AGAGAGCTCACTAAATGCTAATTTAGGCAGAAATAGGAGGTAAAGAAATAGCTAAT-ATC 1471

QY 472 TATTCCCTGAGAGCAGCAGCGGAGGAGCAAGATTCGGATATAAACCCAGGATTCGAGC 531
Db 1472 TATCACCCTGAGAGTACAGGGGAGGAGCAATGATTGGGATAGAAACCCAGGATTCGAGC 1531

QY 532 CGGCAACGGCAACCCCTTTGGTCCCTCTCTTGTATGGCGCTCTGTTTTCACCTCTA 591
Db 1532 CGGCAACGGCAACCCCTTTGGTCTCTCTTGTATGGAGCTCTGTTTTCACCTCTA 1591

QY 592 TT 593
Db 1592 TT 1593

RESULT 28
US-09-997-722-148
; Sequence 148, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 301
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 148
 ; LENGTH: 22436
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-997-722-148

Query Match 76.0%; Score 482.6; DB 11; Length 22436;
 Best Local Similarity 87.6%; Pred. No. 4.8e-139;
 Matches 565; Conservative 0; Mismatches 69; Indels 11; Gaps 3;
 QY 1 CCCTGATCTTTAACTCTCTTTAACTTTGCTCTTCCAGAAATCAAAATGTAAGAACTA 60
 DB 5655 CCTGTACCTTTAACTCTCTTTAACTTTGCTCTTCCAGAAATCGAAGCTGTAAGAACTA 5714
 QY 61 CAAATGTTCTTCAATGAGACCCAGATCGAGTCCATGATCTAAGATCCACCGTGACCC 120
 DB 5715 CAAATGTTCTTCAATGAGACCCAGATCGAGTCCATGATCTAAGATCTACTCGACACC 5774
 QY 121 CTGACCGGCTCTAGCC-CATGCTCCGATGTTAATGACATTAAGAGCAACCCCTCCCGA 179
 DB 5775 CTGACCGGCTCTAGCCCTCATGCTCCATGTTAATGACATTAAGAGCAACCTCTCCCGA 5834
 QY 180 GGAATCTCAACTGCACAAACCTCTATGCTCCCAATTCAGCGGAGCAGTTAGAGCGG 239
 DB 5835 GGAATCTCAACTGCACAAACCTCTATGCTCCCAATTCAGCGGAGCAGTTAGAGCAA 5894
 QY 240 TCATGACCAACCTCCCAACAGACATTTGGTCTTCTGTTGAGAGGGGAGCTGAGAGA 299
 DB 5895 TCATTTGGCCAACTCCCAACAGACATTTGGTCTTCTGTTGAGAGGGGAGCTGAGAG- 5953
 QY 300 CAGGACTAGCTGGATTTCTTAGCCACAGAGATCCCTAAGCTAGCTGGAGGAGTAC 359
 DB 5954 CAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCTAGCTGGAGGAGTAC 6013
 QY 360 TGCAATCACTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 411
 DB 6014 TGCAATCACTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAAGTAGTA 6073
 QY 412 -AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGAGGTTAAGAAATAGCAATCAT 470
 DB 6074 AAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGAGGTTAAGAAATAGCAATCAT 6133
 QY 471 CTATTGCTGAGACACAGCGGAGGACAGAGTCCGGATATAACCCAGGCAATTCGAG 530
 DB 6134 CTATTGCTGAGACACAGCGGAGGACAGAGTCCGGATATAACCCAGGCAATTCGAG 6193
 QY 531 CCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTCTGATGGGCGCTCTGTTTCACTCT 590
 DB 6194 CCGCAATGGCTACCTCTTTGGGTCCCTCCCTTCTGATGGGAGCTCTGTTTACACTCT 6253
 QY 591 ATTTCACCTATTAAATCTTGCAACTGAAATGAAATGAAATGAAATGAAATGAAAT 635
 DB 6254 ATTAACCTTGCAACTGCAAAATATATATATATATATATATATATATATATATAT 6298

RESULT 29

US-10-220-120-15
 ; Sequence 15, Application US/10220120
 ; Publication No. US20040048253A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: PANZER, Scott R.
 ; APPLICANT: SPIRO, Peter A.
 ; APPLICANT: BANVILLE, Steven C.
 ; APPLICANT: SHAH, Purvi
 ; APPLICANT: CHALUP, Michael S.
 ; APPLICANT: CHANG, Simon C.
 ; APPLICANT: CHEN, Alice
 ; APPLICANT: D'SA, Steven A.
 ; APPLICANT: AMSHEY, Stefan
 ; APPLICANT: DAHL, Christopher R.

; APPLICANT: DAM, Tam C.
 ; APPLICANT: DANIELS, Susan E.
 ; APPLICANT: DUFOUR, Gerard E.
 ; APPLICANT: FLORES, Vincent
 ; APPLICANT: FONG, Willy T.
 ; APPLICANT: GREENAWALT, Lila B.
 ; APPLICANT: HILLMAN, Jennifer L.
 ; APPLICANT: JONES, Anissa L.
 ; APPLICANT: LIU, Tommy F.
 ; APPLICANT: ROSEBERRY, Ann M.
 ; APPLICANT: ROSEN, Bruce H.
 ; APPLICANT: RUSSO, Frank D.
 ; APPLICANT: STOCKREHER, Theresa K.
 ; APPLICANT: DAFFO, Abel
 ; APPLICANT: WRIGHT, Rachel J.
 ; APPLICANT: YAP, Pierre E.
 ; APPLICANT: YU, Jimmy Y.
 ; APPLICANT: BRADLEY, Diana L.
 ; APPLICANT: BRATCHER, Shawn R.
 ; APPLICANT: CHEN, Wensheng
 ; APPLICANT: COHEN, Howard J.
 ; APPLICANT: HODGSON, David M.
 ; APPLICANT: LINCOLN, Stephen E.
 ; APPLICANT: JACKSON, Stuart
 ; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: PT-1113 PCT
 ; CURRENT APPLICATION NUMBER: US/10/220,120
 ; CURRENT FILING DATE: 2002-08-26
 ; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,776;
 ; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
 ; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
 ; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
 ; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
 ; 60/204,525; 60/205,285; 60/205,323; 60/205,287;
 ; 60/205,324; 60/205,286
 ; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
 ; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
 ; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
 ; 2000-05-17; 2000-05-17; 2000-05-16; 2000-05-16; 2000-05-15;
 ; 2000-05-16; 2000-05-17; 2000-05-17; 2000-05-17; 2000-05-17;
 ; 2000-05-17; 2000-05-17
 ; NUMBER OF SEQ ID NOS: 422
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 15
 ; TYPE: DNA
 ; LENGTH: 849
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:429446.1:2000FEB18
 US-10-220-120-15

Query Match 75.8%; Score 481.4; DB 17; Length 849;
 Best Local Similarity 89.1%; Pred. No. 2.5e-139;
 Matches 534; Conservative 0; Mismatches 56; Indels 9; Gaps 1;
 QY 27 GTTTGTCTCTCCAGATCAAAATCTGTAAATCTACAAATTTGTTCTTCAAAATGGAGCACCA 86
 DB 121 GTTTGTCTCTCCAGATCAAAATCTGTAAATCTACAAATTTGTTCTTCAAAATGGAGCACCA 180
 QY 87 GATGGAGTCCATGACTAAGATCCACCGTGGACCCCTTGACCGGCTCTAGCCCATGCTC 146
 DB 181 GATGGAGTCTATGACTAAGATCTACACAGACCCCTTGACCGGCTCTAGCCCATGCTC 240
 QY 147 CGATGTTAATGACTTGAAGGACCCCTCCGAGGAAATCTCACTGCAACACCCCTACT 206
 DB 241 CGATGTTAATGACTTGAAGGACCCCTCCGAGGAAATCTCACTGCAACACCCCTACT 300
 QY 207 ATGCCCCAATTCAGCGGAGAGCAGTTAGAGCGGTTCATCAGCAACCTCCCCAACAGCACT 266
 DB 301 ACGCCCCAATTCAGCGGAGAGCAGTTAGAGCGGTTCATCAGCAACCTCCCCAACAGCACT 360

Fri Feb 25 16:26:30 2005

QY 267 TGGGTTTTCTGTTGAGAGGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCCAA 326
DB 361 TGAGTTTTCTGTTGAGAGGGGAGAACTGAGAGACAGACTAGCTGGATTTCTTAGGCCAA 420
QY 327 CGAAGATCCCTAAGCCTAGCTGGGAAGTGCATCCACCTCTAAACATGGGGCTTG 386
DB 421 CTAAGAATCCCTAAGCCTATCTGGGAAGTGCATCCACCTCTAAACATGGGGCTTG 480
QY 387 CAACCTAGCTACACCCGACCAATCA-----GAGAGCTCACTAAATGCTAAATTAG 437
DB 481 CAACCTAGCTACACCCGACCAATCAAGATAGTAAGAGAGCTCACTAAATGCTAAATTAG 540
QY 438 GCAAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCTGAGACAGCGGGAGG 497
DB 541 GCAAAACAGGAGGTAAGAAATAGCCAAATCATCTATTGCTGAGACAGCGGGAGG 600
QY 498 ACAGGATCGGATATTAACCCAGGCACTTCGAGCGGCAACGGCAACCCCTTTTGGGTCC 557
DB 601 ACAATGATCAGGATATTAACCCAGGCACTTCAGCCAGCAATGGCTACCTCTTTGGGTCC 660
QY 558 CTTCCCTTTGATGGGCGCTCTGTTTTCATCTATTTCATCTATAAATCTTGCACCT 616
DB 661 CTTCCGTTTGTATGGGAGCTCTGTTTTCATCTATTTCATCTATAAATCTTGCACCT 719
RESULT 30
US-10-104-047-1148
; Sequence 1148, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1148
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1148
Query Match 73.4%; Score 466; DB 17; Length 2349;
Best Local Similarity 86.4%; Pred. No. 2.5e-134;
Matches 541; Conservative 0; Mismatches 75; Indels 10; Gaps 2;
QY 19 CTTGTTAAGTTTGTCTCTTCAGAACTCAAACTGTAAACTACAAATTTGTTCTTCAAATG 78
DB 1377 CTTCTTCAGTTTGTCTCTTCAGAACTCAAACTGTAAACTACAAATTTGTTCTTCAAATG 1436
QY 79 GAGCACCAGATGGAGTCCATGACTAAGATCCAGTGGACCCCTGGACCGGCTGTAGC 138
DB 1437 GAGCCCCAGATGGAGTCCATGACTAAGATCTACTGCGGACCTCTGGACCGGCTGTAGC 1496
QY 139 CCATGCTCCGATGTTAATGATTTGAAGGACCCCTCCGAGGAAATCTCAACTGACAA 198
DB 1497 CCA-GCTCCCCATTAATGATCAAAAGGACCCCTCCGAGGAAATCTCAACTGATGA 1555
QY 199 CCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGTCAATAGCCAACTCCCCA 258
DB 1556 CACCTACTACGCCCAATTCAGCGGGAAGCAGTTAGAGTGTCTGTGGCAACCTCCCCA 1615
QY 259 ACAGCACTGGGTTTCTGTTGAGAGGGGACTGAGACAGGACTAGCTGGATTTCC 318
DB 1616 ACAGCACTGGGTTTCTGTTGAGAGGGGACTGAGACAGGACTAGCTGGATTTCC 1675
QY 319 TAGGCCAAGAGAAATCCCTAAGCTAGCTGGGAAGTGTGACTGCATCCACCTCTAAACAT 378
DB 1676 TAGGCCAAGAGAAATCCCTAAGCTAGCTGGGAAGTGTGACTGCATCCACCTCTAAACAT 1735

QY 379 GGGGCTTGCAACTTAGCTCACAACCCGACCAATC-----AGAGAGCTCACTAAATG 429
DB 1736 TAGGCTTGCAACTTAGCTCACAACCCGACCAATCAGGTAGTAAGAGAGCTTGGTAAATG 1795
QY 430 CTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTGAGAGCACAG 489
DB 1796 CTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTGAGAGCACAA 1855
QY 490 CGGGGGGACAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCT 549
DB 1856 GGGGCGGGAACATGATCAGGATATAAATCAGGCAATCAAGCCAGCAATGGCTACCACT 1915
QY 550 TTGGGTCCCTCCCTCTGTTTGTATGGGCGCTCTGTTTTCATCTATTTCACCTCTTAAATCT 609
DB 1916 TTGGGTCCCTCCCTCTGTTTGTATGGGAGCTCTGTTTTCATCTATTTCACCTCTTAAATCT 1975
QY 610 TGCAACTGAAAAAAGAAAAAAGAAAAA 635
DB 1976 AAAAAAAGAAAAAAGAAAAA 2001
RESULT 31
US-10-204-887-10
; Sequence 10, Application US/10204887
; Publication No. US20030124569A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; TITLE OF INVENTION: SECRETORY MOLECULES
; FILE REFERENCE: PT-1134 PCT
; CURRENT APPLICATION NUMBER: US/10/204,887
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/185,215; 60/185,216; 60/185,232; 60/205,323; 60/205,287; 60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-05-16; 2000-05-17; 2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1564
; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030124569A1 LI:813218.1:2000FEB01
FEATURE:
NAME/KEY: unsure
LOCATION: 307, 317-318, 335, 918, 921, 943, 1011
OTHER INFORMATION: a, t, c, g, or other
US-10-204-887-10

Query Match 70.7%; Score 449; DB 15; Length 1564;
Best Local Similarity 90.9%; Pred. No. 4.2e-129;
Matches 490; Conservative 0; Mismatches 45; Indels 4; Gaps 1;

QY 41 GAATCAAACTGTAAACTACAAATTTCTTCAAAATGGAGCACCAGATGGAGTCCATGA 100
DB 1022 GAATTGAACTGTAAACTACAAATGTTTCATCAAAATGGAGCCCGCATGATCCATGA 1081

QY 101 CTAAGATCCACCGTGGACCCCTGGACCGGCTGTAGCCCATGCTCCGATGTTAAATGACA 160
DB 1082 CTAAGATCCACCGTAGACCCCGGACCGGTCTCCAGCCCATGCTCTGGTGTAAATGACA 1141

QY 161 TTGAAGSCACCCCTCCGAGGAAATCTCAATGCACAAACCCCTACTATGCCCCCAATTGAG 220
DB 1142 TCGAAGSCACCCCTCCCAAGGAAATCTCAGCTGCACAAACCCCTCTATGCCCCCAGTTGAG 1201

QY 221 CGGAAGCAGTTAGAGGGTTCATCAGCAACCTCCCAACAGCACTGGGTTTTCCTGTT 280
DB 1202 CAGGAAGCAGTTAGAGCAGTTCATCGGCCAACCTCCCAATAGCACTTGGGTTTTCCTGTT 1261

QY 281 GAGAGGGGAGCTGAGACAGGACTAGCTGGATTCTTAGGCCCAACGAAGAATCCCTAA 340
DB 1262 GAGAGTGGGAGCTGAG----AGGACTAGCTGGATTCTTAGGCCGAGTAAAGATCCCTAA 1317

QY 341 GCCTAGCTGGGAAGTGAAGTGCATCCACCTCTAAACATGGGGCTTGCAATTAGCTCACA 400
DB 1318 GCCTAGCTGGGAGGTAACATACATCCATCTTTAAACATGGGGCTTGCAACTAGCTCACA 1377

QY 401 CCGGACCAATCAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAAT 460
DB 1378 CCGGACCAATCAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAAT 1437

QY 461 AGCCAATCATCTATTGCTGAGACAGCGGAGGAGGACAGGATCGGATATAAACCCTA 520
DB 1438 AGCCAATCATCTATTGCTGAGACAGCGGAGGAGGACAGGATCAGGATATAAATCCA 1497

QY 521 GGCATTCCAGCGGCAACGGCAACCCCTTTGGTCCCTCTCCCTTTGTATGGGCGCTCT 579
DB 1498 GGCATTCCAGCGGCAATGGCAACCCCTTTGGTCCCTCTCTTGTATGGGAGCTCTAT 1556

RESULT 32
US-10-220-120-17
Sequence 17, Application US/10220120
Publication No. US20040048253A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: PANZER, Scott R.
APPLICANT: SPIRO, Peter A.
APPLICANT: BANVILLE, Steven C.
APPLICANT: SHAH, Purvi
APPLICANT: CHALUP, Michael S.
APPLICANT: CHANG, Simon C.
APPLICANT: CHEN, Alice
APPLICANT: D'SA, Steven A.
APPLICANT: AMSHEY, Stefan
APPLICANT: DAHL, Christopher R.
APPLICANT: DAM, Tam C.
APPLICANT: DANIELS, Susan E.
APPLICANT: DUFOUR, Gerard E.
APPLICANT: FLORES, Vincent
APPLICANT: FONG, Willy T.
APPLICANT: GREENAWALT, Lila B.

APPLICANT: HILLMAN, Jennifer L.
APPLICANT: JONES, Anissa L.
APPLICANT: LIU, Tommy P.
APPLICANT: ROSEBERRY, Ann M.
APPLICANT: ROSEN, Bruce H.
APPLICANT: RUSSO, Frank D.
APPLICANT: STOCKDREHER, Theresa K.
APPLICANT: DAFFO, Abel
APPLICANT: WRIGHT, Rachel J.
APPLICANT: YAP, Pierre E.
APPLICANT: YU, Jimmy Y.
APPLICANT: BRADLEY, Diana L.
APPLICANT: BRATCHER, Shawn R.
APPLICANT: CHEN, Wensheng
APPLICANT: COHEN, Howard J.
APPLICANT: HODGSON, David M.
APPLICANT: LINCOLN, Stephen E.
APPLICANT: JACKSON, Stuart

TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: PT-1113 PCT
CURRENT APPLICATION NUMBER: US/10/220,120
CURRENT FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;
60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
60/205,324; 60/205,286
PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
2000-05-17; 2000-05-17; 2000-05-17
NUMBER OF SEQ ID NOS: 422
SOFTWARE: PERL Program
SEQ ID NO 17
LENGTH: 1393
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:351965.1:2000FEB01
FEATURE:
NAME/KEY: unsure
LOCATION: 120, 765
OTHER INFORMATION: a, t, c, g, or other
US-10-220-120-17

Query Match 70.3%; Score 446.2; DB 17; Length 1393;
Best Local Similarity 90.3%; Pred. No. 3e-128;
Matches 502; Conservative 0; Mismatches 44; Indels 10; Gaps 2;

QY 47 AAACCTGTAAAACTACAAATTTGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAGA 106
DB 2 AAGCTGTAAACTACTAATTTGTTCTTCAAAAGAGGCCAGATGCACTGACTAAGA 61

QY 107 TCCACCGTGGACCCCTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTCGAG 166
DB 62 TCTACCACAGACCCCTGGACCGGCTGTAGCCCATGCACTGATGTTAATGATTCGANA 121

QY 167 GCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGAA 226
DB 122 GCACCTCC-CCCAAGGAAATTTCACTGCACAAACCCCTACTACACCCCAATTCAGCAGAA 180

QY 227 GCAGTTAGAGCGGTTCATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGG 286
DB 181 GCAGTTAGAGCGGTTCATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGC 240

QY 287 GGGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCCAAGAAATCCCTAAGCCTAG 346

Db	241	GGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCTAAGCCTAG	300
Qy	347	CTGGGAAGGTGACTGCATCCACTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGAC	406
Db	301	CTGGGAAGGTGACCGGGCTACCTTTAAACACGGGGCTTGAACCTTAGCTCACACCCAAC	360
Qy	407	CAATC-----AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGAGGTAAGA	457
Db	361	CAATCAGGTAGTAAAGAGAGCTCACTAAATGCTAAATAGGCAAAACAGAGGTAGAGA	420
Qy	458	AATAGCCAATCATCTATTTCCTGAGAGCACACGGGAGGACAAGATCGGATATAAAC	517
Db	421	AATAGCCAATCATCTATCGCCCTGAGAGCACACGAGGGACAATGATCCGGATATAAAC	480
Qy	518	CCAGGCAATTCGAGCCGGCAACCGCCATCCCTTTGGTCCCTCCCTTTGTAATGGCGCT	577
Db	481	CCNAGCAATTCGAGCCAGCAACGGCTGCGCCCTTTGTGTCCTCCCTCTTGTATGGGAGCT	540
Qy	578	CTGTTTTCACCTATT	593
Db	541	CTGTTTTCACCTATT	556
RESULT 33			
US-10-198-846-9936			
; Sequence 9936, Application US/10198846			
; Publication No. US2003099974A1			
; GENERAL INFORMATION:			
; APPLICANT: Xu, Yongyao			
; APPLICANT: Wang, Youzhen			
; APPLICANT: Steinmann, Kathleen			
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS			
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND			
; TITLE OF INVENTION: THERAPY OF BREAST CANCER			
; FILE REFERENCE: MRI-049			
; CURRENT APPLICATION NUMBER: US/10/198,846			
; CURRENT FILING DATE: 2002-07-18			
; PRIOR APPLICATION NUMBER: 60/306,220			
; PRIOR FILING DATE: 2001-07-18			
; NUMBER OF SEQ ID NOS: 14084			
; SOFTWARE: Fast-Seq for Windows Version 4.0			
; SEQ ID NO 9936			
; LENGTH: 7974			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-198-846-9936			
Query Match 68.4%; Score 434.4; DB 14; Length 7974;			
Best Local Similarity 87.7%; Pred. No. 3.3e-124;			
Matches 537; Conservative 0; Mismatches 56; Indels 19; Gaps 5;			
Qy	1	CCCTGTATCTTTAACTCTTGTAAAGTTGTCTCTCCAGAAATCAAACTGTAAAACTA	60
Db	6409	CCCTGTATCTTTAACTCTTGTAAAGTTGTCTCTCCAGAAATCAAACTGTAAAACTA	6468
Qy	61	CAAAATGTTCTTCAAAATGGAGCCAGATGAGTCCATGAC-TAAGATCCACCGTGGACC	119
Db	6469	CAAAATGTTCTTCAAAATGGAGCCAGATGAGTCCATGAC-TAAGATCTACCATGGACC	6528
Qy	120	CCTGGACGGCTCTGCTAGCCATCTCCGATGTTAATGACATTAAGAGCAACCCCTCCCGA	179
Db	6529	CCTGGACAGGCTGCTAGCCATCTCTGATGTTAATGACATCGAAGACACCCCTCCAGA	6588
Qy	180	GGAATCTCAACTGCAAAACCCCTACTATGCCCCAATTCAGCGGAGAGGATTAGAGCGG	239
Db	6589	GGAATCTCAACTGCAAAACCCCTACTATGCCCCCTGATTACAGAGGAAGTAGTAGTGG	6648
Qy	240	TCATCAGCCACCTCCCCACAGACATTTGGTTTTCTGTTGAGAGGGGGAGCTGAGAGA	299
Db	6649	TTGTGGCCAACTCTCCCAACAGACATCTAGTTTTCTGTTGGAGGGGGAGCTGAGAGA	6708
Qy	300	CAGGACTAGCTGGATTTCTTA-----GGCCAAAGAAATCCCTAAGCCTAGCTGGGA	352
RESULT 34			
US-10-322-281-718			
; Sequence 718, Application US/10322281			
; Publication No. US20040126762A1			
; GENERAL INFORMATION:			
; APPLICANT: David W. Morris			
; APPLICANT: Marc S. Nalandro			
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer			
; FILE REFERENCE: 529452001000			
; CURRENT APPLICATION NUMBER: US/10/322,281			
; CURRENT FILING DATE: 2002-12-17			
; NUMBER OF SEQ ID NOS: 866			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 718			
; LENGTH: 44063			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; NAME/KEY: misc_feature			
; LOCATION: (1)...(44063)			
; OTHER INFORMATION: n = A,T,C or G			
US-10-322-281-718			
Query Match 67.3%; Score 427.6; DB 18; Length 44063;			
Best Local Similarity 86.1%; Pred. No. 9.7e-122;			
Matches 556; Conservative 0; Mismatches 54; Indels 36; Gaps 6;			
Qy	1	CCCTGTATCTTTAACTCTTGTAAAGTTGTCTCTCCAGAAATCAAACTGTAAAACTA	60
Db	1026	CCCTGTATCTTTAACTCTTGTAAAGTTGTCTCTCCAGAAATCAAACTGTAAAACTA	1085
Qy	61	CAAAATGTTCTTCAAAATGGAGCCAGATGAGTCCATGAC-TCAAGTCCCGTGACCC	120
Db	1086	CAAAATGTTCTTCAAAATGGAGCCAGATGAGTCCATGAC-TCTACTGTGACC-	1144
Qy	121	CTGGACCGGCTCTGCTAGCCATCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG	180
Db	1145	-----CCTGCTAGCCCATGCTCCGATGTTAATGACATCGAAGGACACCCCTCCCGAG	1195
Qy	181	GAAATCTCAACTGCAAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAATTTAGACCGT	240
Db	1196	GAAATCTCAACGGCATGATCCCTAGTATGCCCAATTCAGCAGGAAGC-----AGAGCGGC	1251
Qy	241	CATCAGCCAACTCCCAACACACATCTGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	300
Db	1252	CGTGGCCAACTCTCCCAACATCATCTGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	1311
Qy	301	AGGACTAGCTGGATTTCTCTAGGCCCAACGAAGAAATCCCTAAGCCTAGCTGGG-AAAGTGAC	359

Db 1312 AGGACTAACTGGATTTCTTAGGCCAACTAAGAAATCCCAAGCCTAGCTGGGAAAGTGAC 1371
Qy 360 TGCATCCACTCTAAACATGGGGCTTGCACTTAGCTCACCAGCAATC----- 411
Db 1372 TGCACCCACTTTAGACATGGGGCTTTGTAATCAGCTCACCACCGCAATCAGGAGTA 1431
Qy 412 -AGAGAGCTCAGTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAG-CCAATCA 469
Db 1432 AAGAGGGCTCAGTAAATATCAATAGAGCTAAGAGGTAAGAAATAGTCAATCA 1491
Qy 470 TCTATTGCCCTGAGACACAGCGGAGGACAAGGATCGGATATATAACCCAGGCAATCGA 529
Db 1492 TACATTGCCCTGAGACACAGGGGAGGACAATATGATCGGATATAAACCAGGCAATCGA 1551
Qy 530 GCGGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGCGCTCTGTTTTCATC 589
Db 1552 GCAGGAGGGCAACCCCTTTGGGTCCCTCCCTTTTATGGGAGCTCTG----- 1602
Qy 590 TATTTCACCTCTATTAAATCTTGCAACTGAAAAAAGAAAAA 635
Db 1603 -TTTCACCTATTAAATCTGCAACTGAAAAAAGTTAAAA 1646

RESULT 35

US-10-027-632-289680
; Sequence 289680, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 289680
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-289680

Query Match 66.9%; Score 424.6; DB 13; Length 604;
Best Local Similarity 93.4%; Pred. No. 1.1e-121;
Matches 453; Conservative 1; Mismatches 30; Indels 1; Gaps 1;
Qy 110 ACCGTGACCCCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGCA 169
Db 2 ACCATGACCCCTGGACCGGCTGCTAGCAGATGCTCTGATGTTAATGACATCGAGGCA 61
Qy 170 CCCCTCCCGAGGAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCA 229
Db 62 CCCCTCCCGAGGAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCAAGAGCA 121
Qy 230 GTTAGAGCGGTCA-TCAGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGG 288
Db 122 GTTAGAGTGGTCACTCGGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGAGGG 181
Qy 289 GGACTGAGACAGGACTAGTGGATTTCTTAGSCCAAGAAATCCCTAAGCCTAGCT 348

Db 182 GGACTGAGACAGGACTAGTGGATTTCTTAGGCCAACTAAGAAATCCCTAAGCCTAGGT 241
Qy 349 GGGAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACCACCGACCA 408
Db 242 GGGAGGTGACTGCATCCACCTTTAAACACGGGGCTTGCACTTAGCTCACCACCGACCA 301
Qy 409 ATCAGAGAGCTCACTAAAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCCAATC 468
Db 302 ATAAGAGAGCTTACTAAAAATGCTAATTAGGCAAAACAGGAGTAAAGAAATAGCCAATC 361
Qy 469 ATCTATTGGCTGAGACACAGCGGAGGACAAGGATCGGATATATAACCCAGGCAATTCG 528
Db 362 ATCTATTGGCTGAGACACAGTGGGAGGGAATAATGATCGGATATATAACCCAGGCAATTCG 421
Qy 529 AGCCGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGCGCTCTGTTTTCATC 588
Db 422 AGCCAGCAATRGCTACCTCTATGATCCCTCCCTTTGATGGGAGCTCTGTCTTCACT 481
Qy 589 CTATT 593
Db 482 CTATT 486

RESULT 36

US-10-027-632-289680
; Sequence 289680, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 289680
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-289680

Query Match 66.9%; Score 424.6; DB 17; Length 604;
Best Local Similarity 93.4%; Pred. No. 1.1e-121;
Matches 453; Conservative 1; Mismatches 30; Indels 1; Gaps 1;
Qy 110 ACCGTGACCCCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGCA 169
Db 2 ACCATGACCCCTGGACCGGCTGCTAGCAGATGCTCTGATGTTAATGACATCGAGGCA 61
Qy 170 CCCCTCCCGAGGAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCA 229
Db 62 CCCCTCCCGAGGAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCAAGAGCA 121
Qy 230 GTTAGAGCGGTCA-TCAGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGG 288
Db 122 GTTAGAGTGGTCACTCGGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGAGGG 181

```
QY 289 GGACTGAGAGCAGACTAGCTGGATTTCCTAGGCCAACGAGAGATCCCTAAGCTAGCT 348
Db 182 GGACTGAGAGCAGACTAGCTGGATTTCCTAGGCCAACGAGAGATCCCTAAGCTAGCT 241
QY 349 GGGAGAGGTGACTGCTACCTCTTAAACATATGGGGCTTGCACCTTAGCTCACACCCGACCA 408
Db 242 GGGAGAGGTGACTGCTACCTCTTAAACATATGGGGCTTGCACCTTAGCTCACACCCGACCA 301
QY 409 ATCAGAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCAATC 468
Db 302 ATAAGAGAGCTTACTAAATGCTAATAGGCAAAACAGGAGGTAAAGAAATAGCCAATC 361
QY 469 ATCTATTGCTGAGAGCAGCGGAGGACAGAGATCGGATATAAACCCAGGCAATCG 528
Db 362 ATCTATTGCTGAGAGCAGCTGGAGGAGGAAATGATCGGATATAAACCCAGGCAATCG 421
QY 529 AGCCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACT 588
Db 422 AGCCAGCAATRGCTACCTCTATGGATCCCTCCCTTTGTATGGGGCTCTGTTTCACT 481
QY 589 CTATT 593
Db 482 CTATT 486

RESULT 37
US-10-741-600-17699/c
; Sequence 17699, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 7397
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17699
; LENGTH: 50353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17699

Query Match 66.2%; Score 420.4; DB 19; Length 50353;
Best Local Similarity 83.9%; Pred. No. 1.8e-119;
Matches 535; Conservative 2; Mismatches 88; Indels 13; Gaps 5;

QY 11 TTAACCTCCTGTTAAGTTTGTCTCTTCCAGAACTCAAACTGTAAACTACAAATGTTC 70
Db 30181 TTAACCTCCTGTTAAGTTTGTCTCTTCCAGAACTGTAAGCTGTAAACTATAATGTTC 30122
QY 71 TTCAATGGAGCACCAGATGAGTCCATGATCAAGATCCACCGTGGACCCCTGGACCGGC 130
Db 30121 TTCAATGGARCCCAAGATGAGTCTGTGATCAAGATCTACTACAGACCCCTGGACTGGC 30062
QY 131 CTGCTAGCCCATGCTCGATGTTAATGACATTTGAAGCACCCTCCCGAGGAATCTCA 190
Db 30061 CTGCTAGCCCATGCTCCAATGTTGATGACATCGAAAGACCCCTCTCAAGGAATCTCA 30002
QY 191 CTGCACAAACCTTACTATGCCCCAAATTCAGCGGAAGCAGTTAGAGCGGTCAACGCCAA 250
Db 30001 CTGCATGACCCCTTACTATGCCCCAGTTCCGACGAAGCAGTTAGAGCGGTCAACGCCAA 29942
QY 251 CCTCCCCAACAGCACTTGGTTTTTCTGTTTGAAGGGGGGACTGAGAGACAGACTAGCT 310
Db 29941 CCTCCCCAACAGCACTTGGTTTTTCTGTTTGAAGGGGGGACTGTAACAGACAGACTAGCT 29882
QY 311 GGATTTCTTAGGCCAAGAGATCCCTTAAGCTAGCT-GGAGAGTGAAGTGCATCCACC 369
Db 29881 GGATTTCTTAGGCCAATTAAGATTCCTTAAGCTAGCTGGGAGAGGTGACCAACCCACC 29822
QY 370 TCTAAACATGGGGCTTCAACTTAGCTTCACACCCGACCAATC-----AGAGAGCTC 420
```

```
Db 29821 TTTTAAACAAGGGCTAGTAACCTCAGCTCACACCTGACCAATCACATAGTAAGAGAGCTC 29762
QY 421 ACTAAATGCTAATAGGCAAAAA-TAGGAGGTAAAGAAAT-AGCCAATCATCTATTGCC 478
Db 29761 ACTAAATATACCAATTAGGCTTAAAGAGCAGGAGGTAAAGAAATCAATATCTATCACC 29702
QY 479 TGAGAGCACAGCGGAGGACAAAGGATCGGATATATAA-CCAGGCAATTCAGCCGCA 537
Db 29701 TGAGAGCACAGAGGAGGAGGACATGATCGGATATATAA-CCAGGCAATTCAGCCGCA 29642
QY 538 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCAC 597
Db 29641 TGGCAATTCCTTTGGGCGCTCCCATTTATGTTAGTCTGTTTCACTCTATTAAAT 29582
QY 598 TCTATTAAATCTTGCACTGAAAAAAGAAAAA 635
Db 29581 CTTGCACTGCCAAAAAAGAAAAAAGAAAAA 29544

RESULT 38
US-10-466-531-47
; Sequence 47, Application US/10466531
; Publication No. US20040166500A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, Scott R.
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOUR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy F.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAFFO, Abel;
; APPLICANT: MARWAHA, Rakesh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN JR., Edward H.;
; APPLICANT: PERALTA, Careyna H.; DAVID, Marie H.;
; APPLICANT: LEWIS, Samancha A.
; TITLE OF INVENTION: SECRETORY MOLECULES
; FILE REFERENCE: PT-1216 USN
; CURRENT APPLICATION NUMBER: US/10/466,531
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: PCT/US02/01340
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/262,599
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,131
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,164
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,864
; PRIOR FILING DATE: 2001-01-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PERL Program
; SEQ ID NO 47
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: LI:1147914.1:2001JAN12
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 484
; OTHER INFORMATION: a, t, c, g, or other
```

US-10-466-531-47

Query Match 65.4%; Score 415.6; DB 18; Length 1394;
 Best Local Similarity 90.8%; Pred. No. 1.1e-118;
 Matches 532; Conservative 0; Mismatches 44; Indels 10; Gaps 8;

QY 41 GAATCAAACTGTAAACTACAAATTTGTTCTTCAAAATGGAGCACC-AGATGGAGTCCATG 99
 DB 799 GAATTGATCTGTAAACTACAAATTTGTTCTTCAAAATGGAGCACC-AGATGGAGTCCATG 858

QY 100 AC-TAAGATCCACGTGGAGACC-CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATG 157
 DB 859 ACCTAAGATCCACGTGGAGACC-CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATG 918

QY 158 ACATTGAAGCAGCCCTCCGAGGAAATCTCAATGCAACCCCTACTATGCCCAAT 217
 DB 919 ACATCGAAGCAGCCCTCCGAGGAAATCTCAATGCAACCCCTACTATGCCCAAT 978

QY 218 CAGCGGGAAGCAGTTAGAGCGGTGTCATCAGCCAACTCCCAACAGCACTTTGGGTTTCT 277
 DB 979 CAGCAGGAAGCAGTTAGAGCGGTGTCATCAGCCAACTCCCAACAGCACTTTGGGTTTCT 1038

QY 278 GTTGAGAGGGGAGTCTGAGAGCAGTCTAGCTGATTTCTAGGCGCAAGAAATCCC 337
 DB 1039 GTTGAGAGTGGGACTGAGAG-...GAACTAGCTGATTTCTAGGCGCAAGAAATCCC 1095

QY 338 TAAGCCTA-CTGGGAAGGTGACGTGATCCACCTCTAAACATGGGGCTTGCACTTAGC- 395
 DB 1096 TAAGCCTTAGCTGGGAAGGTGACGTGATCCACCTCTAAACATGGGGCTTGCACTTAGC 1155

QY 396 TCACACCCGACCAAT-CAGAGAGCTCACTAAATGCTAAATAGGC-AAAAATAGGAGGTA 453
 DB 1156 TCACACCCGACCAAT-CAGAGAGCTCACTAAATGCTAAATAGGC-AAAAATAGGAGGTA 1215

QY 454 AAGAAATAGCAATCATCTATTGCTGAGAGCAGCGGAGGAGCAAGATCGGATAT 513
 DB 1216 AAGAAATAGCAATCATCTATTGCTGAGAGCAGCGGAGGAGCAAGATCGGATAT 1275

QY 514 AAACCCAGGCAATTCGAGCGGCAAGCAACCCCTTTGGGTCCTCCCTTTGATGGG 573
 DB 1276 AAACCCAGGCAATTCGAGCGGCAAGCAACCCCTTTGGGTCCTCCCTTTGATGGG 1335

QY 574 CGCTCTGTTTCACTCTATTTCATCTATTAAATCTTGCAACTGAA 619
 DB 1336 AGCTCTGTTTCACTCTATTTCATCTATTAAATCTTGCAACCGCA 1381

RESULT 39

US-10-029-386-8554/c
 ; Sequence 8554, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
 ; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 ; FILE REFERENCE: AROMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 8554
 ; LENGTH: 548
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AF045450.1
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.99
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.6

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
 ; OTHER INFORMATION: EST HUMAN HIT: B1087886.1, EVALUATION 1.00e-122
 ; OTHER INFORMATION: EST HUMAN HIT: B1087886.1, EVALUATION 1.00e-122
 ; OTHER INFORMATION: SWISSPROT HIT: P03384, EVALUATION 2.00e-01
 ; OTHER INFORMATION: NT HIT: AL163280.2, EVALUATION 0.00e+00
 US-10-029-386-8554

Query Match 62.3%; Score 395.8; DB 16; Length 548;
 Best Local Similarity 89.5%; Pred. No. 1.1e-112;
 Matches 461; Conservative 0; Mismatches 47; Indels 7; Gaps 3;

QY 1 CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTTCCAGATCAAACTGTAAACTA 60
 DB 513 CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTTCCAGATCAAACTGTAAACTA 454

QY 61 CAAATTTGTTCTTCAAAATGGAGCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
 DB 453 CAAATCAATTTCTTCAAAATGGAGCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 394

QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGGACACCCCTCCCGAG 180
 DB 393 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGGACACCCCTCCCGAG 334

QY 181 GAAATCTCAATCTGCAACACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 240
 DB 333 GAAATCTCAATCTGCAACACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 274

QY 241 CATCAGCCCACTCCCAACAGCAGCTTTGGGTTTCTTGTGAGAGGGGAGCTGAGAGAC 300
 DB 273 CATC-GCCAACTCCCAACAGCAGCTTTGGGTTTCTTGTGAGAGGGGAGCTGAGAGAC 215

QY 301 AGGACTAGCTGGATTTCTTAGGCCAAGCAAGATCCCTAAGCCCTAGCTGGGAAAGGTGACT 360
 DB 214 AGGACTAGCTGGATTTCTTAGGCCAAGCAAGATCCCTAAGCCCTAGCTGGGAAAGGTGACT 155

QY 361 GCATCAGCTCTTAAACATGGGGCTTGCNACTTAGCTCACACCCGACCAATCA-...GAG 415
 DB 154 ACTTCCACCTTTAAACAGGGGGCTTGCNACTTAGCTCACACCCGACCAATCA-...GAG 95

QY 416 AGCTCACTAAATGCTAAATTAGGCAAAATAGGAGGTAAAGATAGCCCAATCATCTATT 475
 DB 94 AGCTCACTAAATGCTAAATTAGGCAAAATAGGAGGTAAAGATAGCCCAATCATCTATT 36

QY 476 GCCTGAGAGCAGCGGAGGAGCAAGGATCGGGA 510
 DB 35 GCCTGAGAGCAGCGGAGGAGCAAGGATCGGGA 1

RESULT 40

US-10-242-355-1175/c
 ; Sequence 1175, Application US/10242355
 ; Publication No. US20030235831A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC003C1
 ; CURRENT APPLICATION NUMBER: US/10/242,355
 ; CURRENT FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: 09/764,897
 ; PRIOR FILING DATE: 2001-01-17
 ; PRIOR APPLICATION NUMBER: 60/179,065
 ; PRIOR FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: 60/180,628
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: 60/214,886
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: 60/217,487
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 60/225,758
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/220,963
 ; PRIOR FILING DATE: 2000-07-26
 ; PRIOR APPLICATION NUMBER: 60/217,496

```

; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1175
; LENGTH: 23855
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-355-1175

Query Match      61.2%; Score 388.4; DB 17; Length 23855;
Best Local Similarity 83.1%; Pred. No. 1.3e-109;
Matches 552; Conservative 0; Mismatches 71; Indels 41; Gaps 8;

Qy   1  CCCTGTATCTTTAAACCTCTTGTAAAGTTTGCTCTCCAGAACTCAAAACGTAAACTA 60
Db   11414 CCCTGTATCTTTAAACCTCTTGTAAAGTTTGCTCTCCAGAACTCGAAAGCTGTAAACTG 11355

Qy   61  CAAATTGTTCTTCAAATGGAGCACAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
Db   11354 CAAACAGTCTTCAAATGACCCCGAGATGCACTCATGACTTAAGACTTACCGGACCC 11295

Qy   121  CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGATTAAGCAATGAAGGCAACCCCTCCCGAG 180
Db   11294 CTGGACGGGTTGCTAGCCATGCTCTGATGTTAATGATCATCGAAGCACCCCTTTCCAAG 11235

Qy   181  GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTGAGCGGGAGCAGATT-AGAGCGG 239
Db   11234 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTGAGCAGGAAGCAGTTAAGAGTGG 11175

Qy   240  TCATCAGCCAACTCTCCCAACAGCACTTGGGTTTTCTGTTCAGAGGGGGGACTGAGAGA 299
Db   11174 TTGTGCGCCAACTTGCACACAGCACTTGGGCTTTCCTGTTGA-CGGGGGAGCTGAGAGA 11116

Qy   300  CAGGACTAGCTGGATTTCCTAGGCCAACGAGAAATCCCTTAAGCCTAGCTGGG-AGGTGA 358
Db   11115 CAGGACTAGCTGGATTTCCTAGGCCGATTAAGAAATTCCTTAAGCCTAGCTGGGAAAGGTGA 11056

Qy   359  CTGCATCCACTCTAAACATGCGGGCTTGCACTTAGCTCACAACCCGACCAATC----- 411
Db   11055 CCGCACCTACTCTTAAACACAGGGGCTTGAATCTAGCTCACAACCCGACCAATCAGGTAGT 10996

Qy   412  --AGAGAGCTCCTAAATAGCTAATTAGGCAAAATAGGAGGTAAAGAAATAG-CCAATC 468
Db   10995 AAAGAGGGCTCGTAAATACAAATTAGGCTAGACAGAAAGGTAAAGACAGATCAATC 10936

Qy   469  ATCTATTGCTGAGAGACACGCGGAGGAGCAAGGATCGGGATATAAAG-CCAGGCATTC 527
Db   10935 ATATATCGCTGAGGGCACAGGGGGAGGACAATGATTGGGATATAAACTCCAGGCATTC 10876

Qy   528  GAGCCGCA-----ACGCAACCCCTTTGGGTCCCTCCCTCTTGTAT 570
Db   10875 GAGCCGGGAGTGGGCAACCTCTTTTCAGGGGGCAACCCCTTTGGGTCCCTCCCTCTTGTAT 10816

Qy   571  GGGGGCTCTGTTTTTCACTCTATTATTCACTCTATTAAATCTTGCAACTGAAAAAATAATA 630
Db   10815 GGGAGCTCTGT-----TTTCACTCTATTAAATCTTGCAACTGCAAAATAATA 10766

Qy   631  AAAA 634
Db   10765 ATAA 10762

```

Search completed: February 21, 2005, 19:18:05
Job time : 365.955 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:52:54 ; Search time 1974.67 Seconds
(without alignments)
12240.400 Million cell updates/sec

Title: US-09-319-156B-6
Perfect score: 635
Sequence: 1 cccgtgatcttttaacctctc.....tgaaaaa.....635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498.8	78.6	1160	3 BC030968	BC030968 Homo sapi
2	492.8	77.6	2748	3 CR605851	CR605851 full-leng
3	488.8	77.0	2749	3 CR617248	CR617248 full-leng
4	488	76.9	2500	3 CR622175	CR622175 full-leng
5	485.6	76.5	771	4 BI087886	BI087886 602852690
6	485	76.4	2748	3 CR625046	CR625046 full-leng
7	470.6	74.1	719	7 CN272394	CN272394 170006000
8	468.6	73.8	1071	5 BX365066	BX365066 BX365066
9	466.6	73.5	758	5 BX357208	BX357208 BX357208
10	464.6	73.2	2716	3 CR613169	CR613169 full-leng
11	454.2	71.5	998	5 BX337769	BX337769 BX337769
12	453.6	71.4	1058	5 BX378303	BX378303 BX378303
13	450.8	71.0	689	9 AG121669	AG121669 Pan.trog
14	437	68.8	494	1 AA781423	AA781423 aj26c03.s
15	435	68.5	1019	5 BX439636	BX439636 BX439636
16	432.4	68.1	522	2 AW971553	AW971553 EST383642
17	430.2	67.7	870	7 CN645411	CN645411 ILLUMIGEN
18	424	66.8	1500	3 BC026287	BC026287 Homo sapi
19	423	66.6	653	9 AG033781	AG033781 Pan.trog
20	415.8	65.5	490	1 A1598135	A1598135 tn4a10.x
21	397.4	62.6	440	2 BE732673	BE732673 601571305
22	396	62.4	543	4 BI963185	BI963185 ie57c10.x
23	395.2	62.2	966	5 BX380176	BX380176 BX380176
24	394.8	62.2	651	9 AG058970	AG058970 Pan.trog

C 25	393.2	61.9	701	9 AG126669	AG126669 Pan.trog
C 26	393	61.9	609	9 AG066901	AG066901 Pan.trog
C 27	392.6	61.8	443	1 AA837267	AA837267 OD26b10.s
C 28	387.2	61.0	712	8 AQ892947	AQ892947 HS_3131.B
C 29	386.8	60.9	436	1 A1128526	A1128526 qc61h10.x
C 30	385.2	60.7	446	1 A1393478	A1393478 tg45g04.x
C 31	384.6	60.6	641	9 AG036829	AG036829 Pan.trog
C 32	384.2	60.5	485	2 AW511366	AW511366 hd45h03.x
C 33	379.8	59.8	664	5 BX481837	BX481837 DXFzp866M
C 34	375.2	59.1	556	1 AUI58595	AUI58595 AUI58595
C 35	369.6	58.2	865	6 CB231128	CB231128 AGENCOURT
C 36	365.8	57.6	470	1 A1074704	A1074704 ox83d05.s
C 37	364.8	57.4	679	9 AG076758	AG076758 Pan.trog
C 38	364.2	57.4	415	1 A1128496	A1128496 qc61e08.x
C 39	360.6	56.8	777	9 AG030228	AG030228 Pan.trog
C 40	359.8	56.7	425	1 A1570707	A1570707 tm79g09.x
C 41	359.8	56.7	431	1 AA552941	AA552941 nk61a10.s
C 42	359.2	56.6	490	1 AA426511	AA426511 zw02e05.r
C 43	359.2	56.6	619	9 AG133542	AG133542 Pan.trog
C 44	359.2	56.6	674	9 AG091649	AG091649 Pan.trog
C 45	355.4	56.0	446	6 CB069106	CB069106 is13a01.x

ALIGNMENTS

BC030968 1160 bp mRNA linear HTC 19-NOV-2003
Homo sapiens cDNA clone IMAGE:4724433, with apparent retained intron.

ACCESSION BC030968
VERSION BC030968.1 GI:22658419
KEYWORDS HTC.
SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 1160)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prance, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, Y.S., Krzywinski, M.I., Skalska, U., Smal, D., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps@mail.nih.gov

Tissue Procurement: CLONTECH

cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 41 Row: m Column: 10
 This clone has the following problem: retained intron.

FEATURES
 source
 1. .1160
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4724433"
 /tissue_type="Placenta"
 /clone_lib="NIH_MGC_79"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"

ORIGIN
 Query Match 78.6%; Score 498.8; DB 3; Length 1160;
 Best Local Similarity 92.6%; Pred. No. 3.6e-133;
 Matches 524; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 70 CTTCAATGGAGCAGCAGATGAGTCCATGACTTAAGATCCACCGTGGACCCCTGGACCGG 129
 Db 582 CTACAAATGGAGCCCAAGATGAGTCCAGACTAGACTTACCGCAGACCCCTGGACCGG 641
 QY 130 CTGTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAGGAATCTCA 189
 Db 642 CTTGCTAGCCCATGCTGATGTTAATGACATCAAAAGGCACCCCTCTGAGGAAATCTCA 701
 QY 190 ACTGCAACACCCCTACTATGCCCCCAATTCAGCGGAGCAGTGTAGACGGTCAATCAGCCA 249
 Db 702 GTGCAACAACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCTGCGCCA 761
 QY 250 ACTCCCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGACAGGACTAGC 309
 Db 762 ACTCCCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGACAGGACTAGC 821
 QY 310 TGGATTTCTAGGCCCAACGAAGATCCCTAAGCTTAGCTTGGGAAGTGTGCTGATCCACC 369
 Db 822 TGGATTTCTAGGCTGACTAAGAATCCCTAAGCTTAGCTTGGGAAGTGTGCTGATCCACC 881
 QY 370 TCTAAACATGGGCTTGCACACTTAGCTCACACCGCAACCAATCAGAGAGCTCACTAAATG 429
 Db 882 TTTAAACACGGGGCTTGCACACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAATG 941
 QY 430 CTAATTTAGGCMAAATAGGAGGTAAAGAAATAGCCCAATCATTTATTCCTGAGAGACAG 489
 Db 942 CTAATTTAGGCMAAAGACAGAGGTAAAGAAATAGCCCAATCATTTATTCCTGAGAGACAG 1001
 QY 490 CGGAGGGACAAGGATCGGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCT 549
 Db 1002 CAGGAGGGACAATGATCGGGATATAAACCCAGTCTTCGAGCGGCAACGGCAACCCCT 1061
 QY 550 TTGGGTCCCTCCCTTGTATGGGCGCTCTGTTTTCACCTTATTTCACTTATTAATCT 609
 Db 1062 TTGGGTCCCTCCCTTGTATGGGCGCTCTGTTTTCACCTTATTTCACTTATTAATCT 1121
 QY 610 TCGAACTGAAAAAATAAAAAA 635
 Db 1122 TCGAACTGAAAAAATAAAAAA 1147

RESULT 2
 CR605851
 LOCUS 2748 bp mRNA linear HTC 21-JUL-2004

DEFINITION

full-length cDNA clone CS0DE012YJ24 of Placenta of Homo sapiens (human).
 ACCESSION CR605851
 VERSION CR605851.1 GI:50486658
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2748)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> InvitroGen Corporation 1600 Faraday Avenue

REFERENCE 2 (bases 1 to 2748)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES
 source
 1. .2748
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE012YJ24"
 /tissue_type="Placenta"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 77.6%; Score 492.8; DB 3; Length 2748;
 Best Local Similarity 90.9%; Pred. No. 2.4e-131;
 Matches 540; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTTTCAGAAATCAAACTGTAAACTA 60
 Db 2167 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTTTCAGAAATCAAACTGTAAACTA 2226
 QY 61 CAAATGTTCTTCAATGGAGCAGACAGATGGATGCATGACTTAAGATCCACCGTGACCC 120
 Db 2227 -----CAAATGGAGCCCAAGATGCAGTCCAAGACTTAAGATCTACCGCAGACCC 2274
 QY 121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGATTAATGAAGGCACCCCTCCCGAG 180
 Db 2275 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGATTAATGAAGGCACCCCTCCCGAG 2334
 QY 181 GAAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240
 Db 2335 GAAATCTCAGCTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 2394
 QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
 Db 2395 CGTGGCCCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 2454
 QY 301 AGGACTAGCTGATTTCTAGSCCAACGAAGATCCCTAAGCTAGCTGGGAAGGTGACT 360
 Db 2455 AGGACTAGCTGATTTCTAGSCCTACTAAGNATCCCTAAGCTAGCTGGGAAGGTGACT 2514
 QY 361 GCATCCACCTCTAAACATGGGGCTTGCACCTTAGCTCAACCCGACCAATCAGAGAGCTC 420
 Db 2515 ACATCCACCTTTAAACACGGGCTTGCACCTTAGCTCAACCCGACCAATCAGAGAGCTC 2574
 QY 421 ACTAAATGCTTAATTTAGGCMAAATAAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCCTG 480
 Db 2575 ACTAAATGCTTAATTTAGGCMAAATAAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCCTG 2634

```

QY 481 AGAGCAGCGGAGGACAAAGGATCGGATATAAAACCCAGGATTCGAGCCCGCAACGG 540
    |||||
Db 2635 AGAGCAGCAGAGGAGGACAATGATCGGATATAAACCCAAAGTCCTTCGAGCCCGCAACGG 2694

QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCTCACTCTATTTT 594
    |||||
Db 2695 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCTCATGCTATTT 2748

RESULT 3
LOCUS CR617248
DEFINITION full-length cDNA clone CS0D1022XJ18 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR617248
VERSION 1 GI:50498055
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2749)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 2749)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
Location/Qualifiers
1. .2749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1022XJ18"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 77.0%; Score 488.8; DB 3; Length 2749;
Best Local Similarity 90.8%; Pred. No. 3.4e-130;
Matches 536; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCCCTTTGTTAAAGTTGTCTCTCCAGAAATCAAAACTGTAATACTA 60
Db 2172 CCCTGTATCTTTAACTCCCTTTGTTAACTTTGTCTCTCCAGAAATCGAAGCTGTAATACTA 2231

QY 61 CAATTTGTTCTTCAATTTGGAGCACCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 2232 -----CAATTTGGAGCACCAGATGAGTCCAGACTAAGATCTACCGCAGACCC 2279

QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 180
Db 2280 CTGGACCGGCTGCTAGCCCATGCTGATGTTTAATGACATTTGAAGCACCCTCCCTCGAG 2339

QY 181 GAAATCTCAACTGCACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 2340 GAAATCTCAGTGCACACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 2399

QY 241 CATCAGCCAACTCCCAACAGCACTGGGGTTTTCTGTTGAGAGGGGGGAGCTAGAGAC 300
Db 2400 CGTCGGCCAACTCCCAACAGCACTAGGTTTTCTGTTGAGATGGGGGAGCTAGAGAC 2459

```

```

QY 301 AGGACTAGCTGATTTCTTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGAAAGGTGACT 360
    |||||
Db 2460 AGGACTAGCTGATTTCTTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGAAAGGTGACT 2519

QY 361 GCATCCACCTCTAAACATGGGCTTGCACACTTAGCTCACACCGGACCAATCACAGAGCTC 420
    |||||
Db 2520 ACATCCACCTTTAAACACCGGCTTGCACACTTAGCTCACACCGGACCAATCACAGAGCTC 2579

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCTG 480
    |||||
Db 2580 ACTAAATGCTAATTAGGCAAAAGACAGAGGTAAGAAATAGCCAAATCATCTATTGCTG 2639

QY 481 AGAGCAGCAGCGGAGGACAAGATCGGATATAAAACCCAGGATTCGAGCCCGCAACGG 540
    |||||
Db 2640 AGAGCAGCAGCGGAGGACAATGATCGGATATAAAACCCAGGATTCGAGCCCGCAACGG 2699

QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCTCACTCT 590
    |||||
Db 2700 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCTATGCT 2749

RESULT 4
LOCUS CR622175
DEFINITION full-length cDNA clone CS0D1051YMI3 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR622175
VERSION 1 GI:50502982
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2500)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 2500)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
Location/Qualifiers
1. .2500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YMI3"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 76.9%; Score 488; DB 3; Length 2500;
Best Local Similarity 91.1%; Pred. No. 5.7e-130;
Matches 534; Conservative 0; Mismatches 40; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCCCTTTGTTAAAGTTGTCTCTCCAGAAATCAAAACTGTAATACTA 60
Db 1925 CCCTGTATCTTTAACTCCCTTTGTTAACTTTGTCTCTCCAGAAATCGAAGCTGTAATACTA 1984

QY 61 CAATTTGTTCTTCAATTTGGAGCACCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 1985 -----CAATTTGGAGCACCAGATGAGTCCAAAGACTAAGATCTACCGCAGACCC 2032

```

QY	121	CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGCAACCCCTCCCGAG	180
Db	2033	CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGCAACCCCTCCCGAG	2092
QY	181	GAATCTCAACTGACCAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240
Db	2093	GAATCTCAGCTGACCAACCTCTACTAGCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	2152
QY	241	CATGACCAACCTCCCAACAGCACTTGGTTTTCTGTGAGAGGGGGACTGAGAGAC	300
Db	2153	CGTCGGCAACCTCCCAACAGCACTTAGTTTTCTGTGAGATGGGGACTGAGAGAC	2212
QY	301	AGGACTAGCTGATTTCTTAGGCCAACGAGAAATCCCTAAGCCTAGCTGGGAGGTGACT	360
Db	2213	AGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAGGTGACC	2272
QY	361	GCATCCACCTCTAAACATGGGCTTGCACATTTAGCTCACACCGCAACCAATCAGAGAGCTC	420
Db	2273	ACATCCACCTTTAAACACGGGCTTGCAATTTAGCTCACACCTGACCAATCAGAGAGCTC	2332
QY	421	ACTAAATGCTAATTAGGCANAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCGCTG	480
Db	2333	ACTAAATGCTAATTAGGCAAGACAGAGAGTAAGAAATAGCCAAATCATCTATTGCGCTG	2392
QY	481	AGAGCACAGCGGAGGACAGGATCGGGATATAAACCCAGGCATTCGAGCCGCAACGG	540
Db	2393	AGAGCACAGCAGGAGGACATGATCGGGATATAAACCCAGGTCTTCGAGCCGCAACGG	2452
QY	541	CAACCCCTTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTTCA	586
Db	2453	CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCA	2498
RESULT 5			
BI087886			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
502852690F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4993894 5',			
mRNA sequence.			
BI087886			
BI087886.1 GI:14506216			
EST.			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1. (bases 1 to 771)			
NIH-MGC http://mgc.nci.nih.gov/ .			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-rc@mail.nih.gov			
Tissue Procurement: ATCC			
cDNA Library Preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: Incyte Genomics, Inc.			
DNA sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L1AM1015 row: d column: 23			
High quality sequence stop: 762.			
Location/Qualifiers			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."			
FEATURES			
source			
1. 771			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:4993894"			
/cell_line="MGC36"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_10"			
/note="Organ: cervix; Vector: pCMV-SPORT6; Site:1			

JOURNAL
Submitted (20-JUL-2004) Genoscope - Centre National de Séquençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1 - Genoscope - Evry - Sequenced with a Next-Clono (dm) primer Five prime
consensus

FEATURES	Location/Qualifiers	source
	1 2748	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_type="taxon.9606"	
	/clone="CS0DI04YK06"	
	/tissue_type="Placenta C	
	/plasmid="pCMVSPORT 6"	

ORIGIN	Query Match	76.4%;	Score 485;	DB 3;	Length 2748;
	Best Local Similarity	91.1%;	Pred. No. 4.3e-129;		
	Matches 531;	Conservative	0;	Mismatches 40;	Indels 12; Gaps 1;
QY	1	CCCTGTATCTTTAACTCCTCTGTGTTAGTTTGTCTCTCCAGAATCAAACTGTAAAACTA	60		
DB	2178	CCCTGTATCTTTAACTCCTCTGTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAAACTA	2237		
QY	61	CAAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCCTGTGACCC	120		
DB	2238	-----CAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC	2285		
QY	121	CTGGACGGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGCACCCTCCCGAG	180		
DB	2286	CTGGACGGGCTGTAGCCACGATCTGATGTTAATGACATCAAAAGCACCCTCTCTGAG	2345		
QY	181	GAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240		
DB	2346	GAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT	2405		
QY	241	CATCAGCCAACTCCCCAAACAGCACTTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	300		
DB	2406	CGTGGGCCAACTCCCCACAGCACTTTAGGTTTTCTCTGTTGAGATGGGGGACTGAGAGAC	2465		
QY	301	AGCACTAGCTGGATTTCTTAGGCCAAACGAGAATCCCTTAAGCCTTAGCTGGGAAGTGACT	360		
DB	2466	AGCACTAGCTGGATTTCTTAGGCTGACTTAAGAATCCCTTAAGCCTTAGCTGGGAAGTGACC	2525		
QY	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420		
DB	2526	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC	2585		
QY	421	ACTAAAATGCTAATTAGGCAAAAATAGGAGGTTAAGAAATAGCCAATCATCTATTGCTGT	480		
DB	2586	ACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAGAAATAGCCAATCATCTATTGCTGT	2645		
QY	481	AGAGCAGCGGGAGGAGCAGAAGGATCGGGATATAAAACCCAGGCAATTCGAGCCGGCAACGG	540		
DB	2646	AGAGCAGCAGGAGGAGCAATGATCGGATATAAACCCCAAGTCTTCGAGCCGGCAACGG	2705		
QY	541	CAACCCCTTTGGGTCCCTCTCTTTGTATGGGGCTCTGTTT	583		
DB	2706	CAACCCCTTTGGGTCCCTCTCTTTGTATGGGAGCTCTGTTT	2748		

RESULT 7	CN272394	719 bp	linear	EST 16-MAY-2004
LOCUS	CN272394			
DEFINITION	17000600005410 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence...			
ACCESSION	CN272394			
VERSION	CN272394.1	GI:47288808		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:			
	Rukarvota: Rukarvota: Rukarvota: Rukarvota: Rukarvota:			

REFERENCE
AUTHORS

TITLE

**JOURNAL
COMMENT**

FEATURES

ORIGIN

Query Match	74.1%	Score 470.6	DB 7	Length 719
Best Local Similarity	92.7%	Pred. No. 4.9e-125		
Matches 494	Conservative 0	Mismatches 35	Indels 0	Gaps 0
QY	75	AATGGACACAGATGGAGTCCATGACTAAGATCCACGTTGAGCCCTCGACCGGCGCTGC	134	
DB	187	AATGGAGCCCAAGATGAGTCCAAAGATTAAGATCTACCGCAGACCCCTGGACCGGCGCTGC	246	
QY	135	TAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAGGAAATCTCAACTGC	194	
DB	247	TAGCCACGATCTGATGTTAATGACATCAAGGCACCCCTCCTGAGGAAATCTCAGCTGC	306	
QY	195	ACAACCCCTACTATGCCCCCAATTTCAGCGGGNAGCAGTTAGAGCGGTCTATCAGCCCACTC	254	
DB	307	ACAACCTCTACTAGCGCCCAATTTCAGCAGGAAGCAGTTTAGAGCGGTCTGTCGCGCAACCTC	366	
QY	255	CCCAACGACCTTGGGTTTTCTCTTCAGAGGGGGGACTGAGACACAGGACTAGCTGGAT	314	
DB	367	CCCAACGACCTTAGGTTTTCTCTTTGAGATGGGGGACTGAGACACAGGACTAGCTGGAT	426	
QY	315	TTCTTAGGCGCAACGAAGAATCCCTTAAGCTAGCTGGGAAGGTGACTGCATCCACTCTAA	374	
DB	427	TTCTTAGGCTGACTAAGATCCCTTAAGCTAGCTGGGAAGGTGACCAATCCACTCTTAA	486	
QY	375	ACATGGGGCTTGCAACTCTTAGCTCACAACCGACCAATCAGAGAGCTCACTAAATGCTAAT	434	
DB	487	ACAGGGGGCTTGCAACTCTTAGCTCACAACGTGACCAATCAGAGAGCTCACTAAATGCTAAT	546	
QY	435	TAGCCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGACACACGCGGA	494	
DB	547	TAGGCAAGACAGGAGGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGACACACGAGGA	606	
QY	495	GGGCAAGGATCGGGATATAAACCCAGGCATTTCGAGCGGGCAACGGCAACCCCTTTGGG	554	
DB	607	GGGCAATGATCGGGATATAAACCCAGGCTCTTCGAGCGGGCAACGGCAACCCCTTTGGG	666	
QY	555	TCCGCTCCCTTTGTATGGGCGCTCTGTTTTCACCTCTATTTCACCTCTATTAAAT	607	
DB	667	TCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACCTCTATTAAAT	719	

RESULT 8
BX365066/c

```

LOCUS      BX365066                1071 bp      mRNA      linear      EST 08-APR-2004
DEFINITION BX365066 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0D1046YA18 3-PRIME, mRNA sequence.
ACCESSION  BX365066
VERSION    BX365066.2  GI:46304105
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1071)
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
COMMENT    On May 5, 2003 this sequence version replaced gi:30374869.

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1A10122C10NP1&c=4215.r.

FEATURES             source
    source
    1..1071
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0D1046YA18"
        /tissue_type="PLACENTA COT 25-NORMALIZED"
        /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT)
        primer. Five prime end enriched, double-strand cDNA was
        digested with Not I and EcoR V
        sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
    Query Match      73.8%; Score 468.6; DB 5; Length 1071;
    Best Local Similarity 89.9%; Pred. No. 2e-124;
    Matches 525; Conservative 3; Mismatches 43; Indels 13; Gaps 2;

    Qy      1  CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGATCAAACTGTAAACTA 60
    Db      571 CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGATCAAACTGTAAACTA 512

    Qy      61  CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
    Db      511 -----CAAATGGAGCCCAAGATGAGTCCAGACTAAGATCTACCCGACACC 464

    Qy      121 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATGAAGGCACCCCTCCCGAG 180
    Db      463 CTGACCGGCTGTAGCCCATGCTCTGATGTTTAATGACATCAAGGCACCCCTCCCTGAG 404

    Qy      181 GAAATCTCAACTGACACACCCCTACTATGCCCCAATTCAGCGGGAACGATGTAGAGCGGT 240
    Db      403 GAAATCTCAGTGCACACACCTCTACTACGCCCAATTCAGCAGGAACGATGTAGAGCGGT 344

    Qy      241 CATCAGCCACCTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGACTGAGAGAC 300
    Db      343 CGTCGGCCAACTCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGACTGAGAGAC 284

    Qy      301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAAGCCTAGCTGGGAAGGTGACT 360
    Db      283 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAAGCCTAGCTGGGAAGGTGACC 224

    Qy      361 GCATCCACCTCTAAACATGGGGCTTGCACTTACTCAGCCCGACCCCAATCAGAGAGCTC 420
    Db      223 ACATCCACCTTTAAACAGGGGGCTTGCACTTACTCAGCCCGACCCCAATCAGAGAGCTC 164

LOCUS      BX357208                758 bp      mRNA      linear      EST 08-APR-2004
DEFINITION BX357208 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0D1022YJ18 3-PRIME, mRNA sequence.
ACCESSION  BX357208
VERSION    BX357208.2  GI:46305595
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 758)
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
COMMENT    On May 5, 2003 this sequence version replaced gi:30376125.

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1022DE09NP1&c=4215.r.

FEATURES             source
    source
    1..758
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0D1022YJ18"
        /tissue_type="PLACENTA COT 25-NORMALIZED"
        /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT)
        primer. Five prime end enriched, double-strand cDNA was
        digested with Not I and EcoR V
        sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
    Query Match      73.5%; Score 466.6; DB 5; Length 758;
    Best Local Similarity 88.8%; Pred. No. 7.1e-124;
    Matches 509; Conservative 9; Mismatches 43; Indels 12; Gaps 1;

    Qy      1  CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGATCAAACTGTAAACTA 60
    Db      576 CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGATCAAACTGTAAACTA 517

    Qy      61  CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
    Db      516 -----CAAATGGAGCCCAAGATGAGTCCAGACTAAGATCTACCCGACACC 469

    Qy      121 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATGAAGGCACCCCTCCCGAG 180
    Db      468 CTGACCGGCTGTAGCCCATGCTCTGATGTTTAATGACATCAAGGCACCCCTCTCTGAG 409

    Qy      181 GAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCAGCGGAAGCAGTGTAGAGCGGT 240
```

```

Db      408  GAAATCTCAGCTGCAACACCTCTACTACGCCCAATTCACGAGGAGCAGTGTAGAGCGGT 349
Qy      241  CATCAGCAACCTCCCAACAGCACTGGGTTTCTGTTTGGAGGGGGGACTGAGAGAC 300
Db      348  CGTCGGCAACCTCCCAACAGCACTTAGGTTTCTGTTTGGAGATGGGGGACTGAGAGAC 289
Qy      301  AGGACTAGCTGGATTTCCTAGGCCAACGAGAGATCCCTAAGCCTAGCTGGGAAAGGTGACT 360
Db      288  AGGACTAGCTGGATTTCCTAGGCCAACGAGAGATCCCTAAGCCTAGCTGGGAAAGGTGACC 229
Qy      361  GCATCCACCTCTAAACATGGGGTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db      228  ACATCCACCTTTAAACAGCGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 169
Qy      421  ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGGCTG 480
Db      168  ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTSCYS 109
Qy      481  AGACACAGCGGGAGGACAGGATCGGGATATAAACCCAGGCATTCGAGCGGCACGG 540
Db      108  AGACACAGCAGGAGGAGCAATGATCGGGATATAAACCCAGGCATTCGAGCGGCACGG 49
Qy      541  CAACCCCTTTGGTCCCTCCCTTTCTGATGGG 573
Db      48  CMACCCCTTTGGTCCCTCCCTTTGTAATGGG 16

```

```

RESULT 10
LOCUS   CR613169          2716 bp    mRNA      linear      HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODE013Y120 of Placenta of Homo sapiens
(human).
ACCESSION CR613169.1 GI:50493976
VERSION   HTC; CNSLT_CDNA.
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 2716)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished
REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
2 (bases 1 to 2716)
REFERENCE Genoscope.
AUTHORS   Direct Submission
TITLE     Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL   BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT   1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

```

```

FEATURES
source
1..2716
/mol_type="mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODE013Y120"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

```

```

ORIGIN
Query Match 73.2%; Score 464.6; DB 3; Length 2716;
Best Local Similarity 90.9%; Pred. No. 3.6e-123;
Matches 510; Conservative 0; Mismatches 39; Indels 12; Gaps 1;
Qy 1 CCCTGTATCTTTAAACCTCCTGTTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60

```

```

Db      2168  CCTGTATCTTTAAACCTCCTGTTTAACTTTGTCTCTTCCAGATCGAAGCTGTAAACTA 2227
Qy      61  CAAATTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAAGATCCACCGTGGACCC 120
Db      2228  -----CAAATGGAGCCCAAGATGCGAGTCCAAAGACTAAGATCTACCGAGACCC 2275
Qy      121  CTGACCGCGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCAACCCCTCCGAG 180
Db      2276  CTGACCGCGCTCTAGCCCATGCTGATGTTTAATGACATCAAAAGGCACCCCTCCTGAG 2335
Qy      181  GAAATCTCAATGTCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGAGCTTTAGAGCGGT 240
Db      2336  GAAATCTCAGCTGCAACAACCTCTACTACGCCCAATTCAGCAGGAAAGCAGTTAGAGCGGT 2395
Qy      241  CATCAGCCCAACCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db      2396  CGTCGGCCCAACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGAGCTGAGAGAC 2455
Qy      301  AGGACTAGCTGGATTTCCTAGGCCAACGAGATCCCTAAGCCTAGCTGGGAAAGGTGACT 360
Db      2456  AGGACTAGCTGGATTTCCTAGGCCGACTAAGATCCCTAAGCCTAGCTGGGAAAGGTGACC 2515
Qy      361  GCATCCACCTCTAAACATGGGGCTTGCATTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db      2516  ACATCCACCTTTAAACACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 2575
Qy      421  ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGGCTG 480
Db      2576  ACTAAATGCTAATTAGGCAAAAGCAGGAGGTAAAGAAATAGCAATCATCTATTGGCTG 2635
Qy      481  AGAGCAGCGGGAGGAGGACAAAGGATCGGGATATAAACCCAGGCATTCGAGCGGCACGG 540
Db      2636  AGAGCAGCAGGAGGAGGACAAATGATGGGATATAAACCCAGGCTTCGAGCGGCACGG 2695
Qy      541  CAACCCCTTTGGTCCCTCCCTC 561
Db      2696  CAACCCCTTTGGTCCCTC 2716

```

```

RESULT 11
LOCUS   BX337769/c      998 bp    mRNA      linear      EST 07-APR-2004
DEFINITION BX337769 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI051Y13 3-PRIME, mRNA sequence.
ACCESSION BX337769
VERSION   BX337769.2 GI:46272079
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 998)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   On May 2, 2003 this sequence version replaced gi:30337641.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r

```

```

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CSODI051Y13&c=4215.r.
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
FEATURES
source
1..998

```



```

DEFINITION Pan troglodytes DNA, clone: PTB-130M15.F, genomic survey sequence.
ACCESSION AG121669
VERSION AG121669.1 GI:16650834
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totohi,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 689)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totohi,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimps@gsc.riken.go.jp URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .689
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-130M15.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
Query Match 71.0%; Score 450.8; DB 9; Length 689;
Best Local Similarity 88.3%; Pred. No. 2.7e-119;
Matches 515; Conservative 0; Mismatches 58; Indels 10; Gaps 2;

Qy 1 CCTGTATCTTTAACTCTTGTGTTAGTTTGTCTTCCAGATCAAACTGTAATACTA 60
Db 666 CCNTGGATCTTTAACTCTTGTGTTAGTTTGTCTTCCAGATCAAACTGTAATACTA 607

Qy 61 CAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 606 TAATAGTTTCTTCAATGGAGCACCCTTATGCAGTCCATGACTAAGATCTACCATGGACCC 547

Qy 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTGAAGCACCCTCCCGAG 180
Db 546 CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTGAAGCACCCTCCCGAG 487

Qy 181 GAAATCTCAATGTGAC -AACCCCTACTATGCCCCAATTCAGCGGGAAGAGTAGACGG 239
Db 486 GAAATCTCAATGTGACCAAAACCTTACTACACTCCAGTTCAGCAGGAAGAGTAGACGG 427

Qy 240 TCATCAGCAACCTCCCAACAGACATTTGGTTTCTCTGTTGAGAGGGGGAGCTGAGAGA 299
Db 426 TCGTCGGCAACCTCCCAATGGCACTTGGGTTTCTCTGTTGAGAGGGGGAGCTGAGAGA 367

Qy 300 CAGGACTAGCTGGATTCTTAGGCCAACAGAGAAATCCCTAAGCCCTAGCTGGGAAGGTGAC 359
Db 366 CAGGACTATCTGGATTCTTAGGCCAACAGAAATCCCTAAGCCCTAGCTGGGAAGGTGAC 307

Qy 360 TGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACAACCCGACCAATC----- 411
Db 306 CACATTTCATTTTAAACACAGGGCTTGCAACTTAGCTCACAACCCGACCAATCAGGTAGTA 247

```

```

Qy 412 -AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCAT 470
Db 246 AGAGGGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCAT 187

Qy 471 CTATTGCTTGAGAGACAGCGGAGGAGGACAGGATCGGATATTAACCCAGGATTCGAG 530
Db 186 TTATTGCTTGAGAGTACAGCGGAGGAGGACAGGATCGGATATTAACCTCATGCAATTCGAG 127

Qy 531 CCGGCAACGCGCAACCCCTTTGGGTCCCTCCCTTGTATGGG 573
Db 126 CCACCAATGCTACCTCTTTGGGTCCCTCCCTTGTATGG 84

RESULT 14
LOCUS AA781423/c
DEFINITION aa78c03.s1 Soares testis_NHT Homo sapiens cDNA clone 1391428 3',
similar to contains PTR7.tl PTR7 repetitive element ;, mRNA
sequence.
ACCESSION AA781423
VERSION AA781423.1 GI:2840754
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 494)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldino, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 1645 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 475.
FEATURES
source
1. 494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="1391428"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis_NHT"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [s,
TGTTACCAATCTGAGTGGAGCGGCGCCCAATTTTGTATTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldino."

ORIGIN
Query Match 68.0%; Score 437; DB 1; Length 494;
Best Local Similarity 92.7%; Pred. No. 2.5e-115;
Matches 456; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 131 CTGCTAGCCCATCTCGATGTTAATGACATTAAGCACCCTCCCGAGAAATCTCAA 190
Db 494 CTGCTAGCCCATCTCGATGTTAATGACATTAAGCACCCTCCCGAGAAATCTCAA 435

```

double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN	Query Match	Score	DB 5	Length	1019
	Best Local Similarity	83.4%	Pred. No. 1.1e-114		
	Matches	494	Conservative	19	Mismatches 66; Indels 13; Gaps 2;
Qy	1	CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAACAACTA	60		
Db	579	CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAACAACTA	520		
Qy	61	CAATGTTCTTCAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC	120		
Db	519	-----CAATGGAGCCCCAAGATGAGTCCATGACTAAGATCCACCGTGGACCC	472		
Qy	121	CTGACCGGCTGTAGCCCATGCTCGATGTTTAATGACATTAAGAGCAGCCCTCCGAG	180		
Db	471	CTGACCGGCTGTAGCCCATGCTCGATGTTTAATGACATTAAGAGCAGCCCTCCGAG	412		
Qy	181	GAATCTCACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240		
Db	411	GAATCTCACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	352		
Qy	241	CATCAGCAACCTCCCCCAACAGCACCTTTGGGTTTTTCTGTTGAGAGGGGGAAGCAGT	300		
Db	351	CGTCGGCAACCTCCCCCAACAGCACCTTTGGGTTTTTCTGTTGAGAGGGGGAAGCAGT	292		
Qy	301	AGGACTAGTGGATTTCTTAGGCCAAGCAAGATCCCTAAGCCTAGCTGGGAGGTGACT	360		
Db	291	AGGACTAGTGGATTTCTTAGGCCAAGCAAGATCCCTAAGCCTAGCTGGGAGGTGACT	232		
Qy	361	GCATCCACCTTAACATGGGCTTGCAACTTACTCACCACCCCAACCAATCAGAGAGCTC	420		
Db	231	ACATCCACCTTTAAACAGGGGCTTGCAACTTACTCACCACCCCAACCAATCAGAGAGCTC	172		
Qy	421	ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG	480		
Db	171	ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG	112		
Qy	481	AGAGCAGCGGGAGGACAGGATCGGATATAACCCAGGCAATTCGAGCGGCAACGG	540		
Db	111	AGAGCAGCAGGAGGAGGACATATCGGGATATAAAYCCAAAGTCTTCGAGCGGCAACAG	52		
Qy	541	CAACCCCTTTGGGTCCTCCCTTTGATGGGCTCTGTTTTCACCTAT 592			
Db	51	AAACCCYTTT-TTTCYCCCTTTTGTGNDTTTTTTTTTNTTTTATTT 1			

RESULT 16
 AW971553/c
 LOCUS
 DEFINITION
 EST383842
 AW971553
 AW971553.1
 GI:8161399
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 522)
 Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
 Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and
 Quackenbush, J.
 Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)
 Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208

Qy	191	CTGCACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTTCATCAGCCAA	250
Db	434	CTGCACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTTCATCAGCCAA	375
Qy	251	CCTCCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGGAGCTGAGACAGGACTAGCT	310
Db	374	CCTCCCCAACAGCACTTGGGTTTTCTGTGAGATGGGGAGCTGAGACAGGACTAGCT	315
Qy	311	GGATTTCTTAGCCACAGGAATCCCTAAGCCTAGCTGGGAAGTGAAGTGCATCCACCT	370
Db	314	GGATTTCTTAGCCACAGGAATCCCTAAGCCTAGCTGGGAAGTGAAGTGCATCCACCT	255
Qy	371	CTAAACATGGGCTTGCACCTTAGCTCACACCCGACCAATCAGAGAGTCACTAAATGC	430
Db	254	TTAAACAGGGCTTGCACCTTAGCTCACACCTGACCAATCAGAGAGTCACTAAATGC	195
Qy	431	TAAATTAGGCAAAATAGGAGGTAAGAAATAGCCATCATCTATTGGCTGAGAGCACAGC	490
Db	194	TAAATTAGGCAAAATAGGAGGTAAGAAATAGCCATCATCTATTGGCTGAGAGCACAGC	135
Qy	491	GGGAGGACAGGATCGGATATAAACCCAGGCATTCGAGCGGCAACGGCAACCCCTT	550
Db	134	AGGAGGACAGGATCGGATATAAACCCAGGCATTCGAGCGGCAACGGCAACCCCTT	75
Qy	551	TGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTCACTTAAATCTT	610
Db	74	TGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTCACTTAAATCTT	15
Qy	611	GCAACTGCAAAAAA 624	
Db	14	GCAACTGCAAAAAA 1	

RESULT 15
 BX439636/c
 LOCUS
 DEFINITION
 BX439636 Homo sapiens PLACENTA Homo sapiens cDNA clone CSODE012VJ24
 3-PRIME, mRNA sequence.
 ACCESSION
 BX439636
 VERSION
 BX439636.2
 GI:47000005
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1019)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 15, 2003 this sequence version replaced gi:30771765.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CSODE012DE12NP1&c=4215.r.
 Location/Qualifiers

FEATURES
 source
 1. .1019
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODE012VJ24"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="vector: pCMVSPORT_6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,

Email: johnq@tigr.org
Plate: 292
Seq primer: Forward

Location/Qualifiers
1. 522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="WAGE ressequencing, MAGL"
/note="vector: pBluescriptSKm"

FEATURES

Source

ORIGIN

Query Match 68.1%; Score 432.4; DB 2; Length 522;
Best Local Similarity 91.8%; Pred. No. 5.5e-114;
Matches 479; Conservative 0; Mismatches 41; Indels 2; Gaps 2;
Qy 97 ATGACTAAGATCCACCGTGAGACCCCT-GGACCGGCTGTAGCCCATGCTCCGATGTAA 155
Db 522 AGGTTTAAGATCTTCGCGGACCCCTAAACCTGGCTTGCTAGGCCATGTTCTGGTGTAA 463
Qy 156 TGACATTTGAGGACCCCTCCCGAGGAA-TCTCAACTGCACACCCCTACTATGCCCA 214
Db 462 TGACATCGAAGTCACTCCTCTGAGGAAATTTCTCAACTGCACACCCCTTATTATGCCCA 403
Qy 215 ATTGAGCGGGAAGCAGTTAGAGCGGTCTAGCCCACTCCCAACAGCACTTGGGTTT 274
Db 402 GTTCAGGAGGAGGAGTTAGGTGCTCATCGCCAACTCACCACAGCACTTGGGTTT 343
Qy 275 CCTGTTGAGGCGGGAAGTCTGAGACAGGACTAGCTGGATTTCTAGGCCAACGAGAA 334
Db 342 CCTGTTGAGAGGAGGAGTCTGAGACAGGACTAGCTGGATTTCTAGGCCGATTAAGA 283
Qy 335 CCCTAAGCTAGTGGGAGGTGACTGCATCCACCTCTAACATGGGGCTTGCAACTAG 394
Db 282 CCCTAAGCTAGTGGGAGGTGACCGCATCCACCTTTAAACACGCGGCTTGCAACTAG 223
Qy 395 CTCACCCGACCAATCAGAGAGTCTACTAAATGCTAATTAGGCAAAAATAGAGGTAA 454
Db 222 CTCACCCGACCAATCAGAGAGTCTACTAAATGCTAATTAGGCAAAAATAGAGGTAA 163
Qy 455 AGAATAGCCCAATCATCTATTGCTGAGAGCAGCGGGAGGGAACAAGGATCGGATATA 514
Db 162 AGAATAGCCCAATCATCTATTGCTGAGAGCAGAGTGGGAGGGAACAAGGATTCGA 103
Qy 515 AACCCAGCATTCAGCGGCAACGGCAACCCCTTTGGGTGCTCCCTTCTTGTATGGC 574
Db 102 AACCCAGCATTCAGCGGCAACGGCAACCCCTTTGGGTGCTCCCTTCTTGTATGGG 43
Qy 575 GCTGTTTTTCACTCTATTTTCACTCTATTAAATCTTGCAACT 616
Db 42 GCTGTTTTTCACTCTATTTTCACTCTATTAAATCTTGCAACT 1

RESULT 17

CN645411 870 bp mRNA linear EST 13-MAY-2004
LOCUS ILLUMIGEN_MCO_23968 Katze_MMSP Macaca mulatta cDNA clone
DEFINITION IBUW:10227 5' similar to Bases 165 to 770 highly similar to human
Unigene Hs.349001, mRNA sequence.
CN645411
CN645411.1 GI:47158854
EST.
Macaca mulatta (rhesus monkey)
SOURCE Macaca mulatta
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
1 (bases 1 to 870)
Katze.M.G.; Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
Large-scale Rhesus Macaque cDNA Sequencing
Unpublished (2003)
Contact: C. Magness
Illumigen Biosciences Inc.

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.02.24. 676 Q20 bases.
PCR Primers
FORWARD: CCTCTACTAAGGGAACAAAA
BACKWARD: CACTATAGGCGCAATTGGTA
Insert Length: 870 Std Error: 0.00
Plate: CL000135 row: H column: 02
Seq primer: CCTCTACTAAGGGAACAAAA
POLYA=No.

FEATURES

Source

Location/Qualifiers

1. 870
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBUW:10227"
/sex="male"
/cell_type="mononuclear lymphocyte"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_lib="Katze_MMSP"
/notes="Organ: spleen; Vector: Uni-ZAP XR; Site: 1: EcoR I;
Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis
kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold
Cloning kit (Catalog #200450)"

ORIGIN

Query Match 67.7%; Score 430.2; DB 7; Length 870;
Best Local Similarity 83.8%; Pred. No. 2.6e-113;
Matches 527; Conservative 0; Mismatches 88; Indels 14; Gaps 3;
Qy 21 TGTAAAGTTGTCTCTCCAGAAATCAAACTGTAAACTACAAATCTTCTCAAAATGGA 80
Db 166 TTTTAAATTTGTTTCTTCCAGAAATCGAAAGCTGTAAAACTACAAATGTTCTTCAAAATGGA 225
Qy 81 GCACCATGAGTCCATGATCAATGATCCACCGTGGACCCCTGGACCGGCTGTAGGCC 140
Db 226 GCCTCCGATCGATCCATGATCAATGATCTACTGCGGACCCCTGGACCGCTGTATATCC 285
Qy 141 ATGCTCCGATGTTAATGACATTAAGGACACCCCTCCGAGGAAATCTCAACTGCACAAAC 200
Db 286 ATGCTCTGATGTTGGTGACATCGAAGGACCCCTCCGAGGAAATCTCAACTGCATGAC 345
Qy 201 CCTACTATGCCCAATTCAGGGGAGCAGTGTAGCGGTCATCAGCCAACTCCCAAC 260
Db 346 CCTACTACGCTCAGTTTCAGCAGGAGCAGTGTAGCAGCATCAGCCAACTCCTCAAC 405
Qy 261 AGCACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTA 320
Db 406 AGCACTTGGATTTTCTGTTGTAAGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTA 465
Qy 321 GGCCAAACGAAATCCCTAAGCCCTAGCTGGGAAGGTGACTGCATCTCACTCTAAACATGG 380
Db 466 AGCCAACTAAGATCCCTAAGCCCTAGCTGGGAAGGTGACTGCATCTCTTTAAACACGG 525
Qy 381 GGCTTGCACTTAGCTCAGACCCGACCAATC-----AGAGGCTCACTTAAATGCT 431
Db 526 GGCTGCAATTTAGCTCAGACCCGATCAATCAGGTAGTAAAGAGAGCTCACTTAAATGTT 585
Qy 432 AATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATC--TATTGCTCGAGACACAG 489
Db 586 AATTAAAGCAAAAACAGGAGGTAAAGAAATAGTCAATCACCTTATCGCTGAGAGCAG 645
Qy 490 CGGAGGAGCAGAGGATCGGATATAAACCCAGGCAATTCGAGCCCGGCAACGGCAACCCCT 549
Db 646 GAGGAGGACCAATGATAGGAGATATAAACCCAGGCAATTCAGAGTCAAGCAATGGCAACCCCT 705
Qy 550 TT---GGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTCTTCTCTTCTTATAA 606
Db 706 TTTTGGTCCCTCCCTCCCAATTTTATGGGAGCTCTGGGTTTCTCTTATAAATCTTGCAT 765


```

Qy 607 TCTTCCAACTGAAAAAATAAAAAA 635
Db 766 GCTCAAAAAACATATTAAAAATAAAA 794

RESULT 18
BC026287 1500 bp mRNA linear HTC 04-MAR-2003
LOCUS Homo sapiens, clone IMAGE:4770655, mRNA.
DEFINITION BC026287
ACCESSION BC026287.1 GI:22382161
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (02-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 37 Row: m Column: 12
This clone has the following problem: retained intron.
FEATURES
    Location/Qualifiers
        1..1500
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4770655"
            /tissue_type="Placenta"
            /clone_lib="NIH MGC 79"
            /lab_host="DH10B"
            /note="Vector: pDNR-LIB"
ORIGIN
Query Match 66.8%; Score 424; DB 3; Length 1500;
Best Local Similarity 85.7%; Pred. No. 1.9e-111;
Matches 557; Conservative 0; Mismatches 75; Indels 18; Gaps 7;

Qy 1 CCTGTATCTTTAACTCTCTTGAAGTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
Db 849 CCCTGTATCTTTAACTCTCTTGAAGTTGTCTCTTCAGAAATCAAAAGCTGTAAACAA 908

Qy 61 CAAATTGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 909 CACATCGTTCTTCAAAATGGAGCACCAGATGCATGCATGATCAATCAACCGGATCC 968

Qy 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGAGCAACCCCTCCCGAG 180
Db 969 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGAGCAACCCCTCCCGAG 1028

Qy 181 GAAATCTCACTGCACAAACCCCTACTATGCCCCAAATTCAGGGGAGCAGTTAGCGGT 240
Db 1029 GAAATCTCAATGCACAAACCCCTACTATGCCCCAAATTAAGCAGGAAGCAGTTGGAGCAGT 1088

```

```

Qy 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGGACTGAGAGAC 300
Db 1089 CGACGGCCAACTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGTACTGAGAGAC 1148

Qy 301 AGGACTAGCTGGATTTCCTAGGCAACGAGAGATCCCTAAGCCCTAGCT-GGGAAGGTGAC 359
Db 1149 AGGACTAGCTGGATTTCCTAGGCGGACTAAACAATTTCTAAGCTTTAGTGGGGAAGGTGAC 1208

Qy 360 TGCATCCACCTCTAAACATGGGGC--TTGCAACTTAGCTCACACCCGACCAATC----- 411
Db 1209 TGCATCCACCTTTAAACACAGGACTTTTGTAACTCAGCTCACATCTGCGCAATCAGGTAG 1268

Qy 412 ---AGAGAGCTCACTAAATGCTAATPAGGC-AAAATAGAGGTAAGAATAAGCCAAT 467
Db 1269 TAAAGAGAGCTCATTTAAATACCAATTAGTCTAAAAACAGGAGGTAAGAATAA---AAT 1325

Qy 468 CATCTATTGCTGAGACACAGCGGAGGACAGGATCGGATATATAA-CCCAAGGCATT 526
Db 1326 CATCTCTCGCTGACAGCAGAGGAAGGGGCAATGATGGATATATAACCCCAAGGCATT 1385

Qy 527 CGAGCCGGCA-ACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTGTTTC 585
Db 1386 CGAGCTGGAGTGGCAACCCCTTTGGGTCCCTCCCTCCCTTTGTATGGAGCTCTGTGTTTC 1445

Qy 586 ACTCTATTCTACTCTATTAATCTTGCACTTGCACTGAAAAAATAAAAAA 635
Db 1446 ACTCTATTCAATCTTGCACTTGCACTGAAAAAATAAAAAAATAAAAAA 1495

RESULT 19
AG033781/c 653 bp DNA linear GSS 01-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-008021.F, genomic survey sequence.
DEFINITION AG033781
ACCESSION AG033781
VERSION AG033781.1 GI:16560654
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
JOURNAL Totoki, Y., Watanabe, H. and Sakaki, Y.
REFERENCE BAC end sequences of Library PTB
AUTHORS Unpublished
TITLE 2 (bases 1 to 653)
JOURNAL Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
AUTHORS Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
    Location/Qualifiers
        1..653
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-008021.F"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

```


Query Match 66.6%; Score 423; DB 9; Length 653;
 Best Local Similarity 88.0%; Pred. No. 3e-111;
 Matches 51; Conservative 0; Mismatches 52; Indels 18; Gaps 4;
 2 CCTGTATCTTAACTCC-TTGTAAAGTTGTCTTCAGAAATCAAACTGTAAACTA 60
 639 CTGTGTTCTTAAACCCCGTGTAAAGTGTCTTCCAGAAATCANAGCTGTAAAGCTA 580
 61 CAAATGTTCTTCAAAATGGAGCAGCAGTGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 579 CAAATGTTCTTCAAAATGGAGCAGCAGTGGAGTCCATGACTAAGATCCACCGTGGACCC 535
 121 CTGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
 534 CTGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 476
 181 GAAATCTCAACTGACACACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 240
 475 AAAATCTCAACTGACACACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTTAGAGCGGT 416
 241 CATCAGCAACCTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCAGTGGAGGAC 300
 415 CGTTGGCCAACTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCAGTGGAGGAC 356
 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCCTAGCTGGGAAAGTGACT 360
 355 AGGACTAGCTGGATTTCTTAGGCCCAACGAAAGATCCCTAAGCCTAGCTGGGAAAGTGACT 296
 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCA-GAGAGCT 419
 295 ACATCCACCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCA-GAGAGCT 236
 420 CACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCCT 479
 235 CACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCCT 176
 480 GAGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCGGCAACG 539
 175 GAGAGCACAGGAGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCGGCAACG 116
 540 GCAACCCCTTTTGGGTCCTCCCTCTTGTATGGCGCTCTG 580
 115 GCTACCTCTTTGGGTCCTCCCTCTTGTATGGAGCTCG 75

RESULT 20
 AI598135/c
 LOCUS
 DEFINITION
 tnl410.x1 NCI CGAP Brn25 Homo sapiens cDNA clone IMAGE:2167578 3', similar to contains_PFR5.tl PFR5 repetitive element ;, mRNA sequence.
 AI598135
 AI598135.1 GI:4607183
 EST.
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 490)
 NCBI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGA), Tumor Gene Index
 Unpublished (1998)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1407 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 455
 POLYA=No.

FEATURES

source

1..490
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2167578"
 /tissue_type="anaplastic oligodendroglioma"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Brn25"
 /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I oligo(dT) primer [5' TGTTACCAATCTGAATGGAGCGCGCATAGGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

	Query Match	65.5%;	Score 415.8;	DB 1;	Length 490;
	Best Local Similarity	92.9%;	Pred. No. 3.5e-109;		
	Matches 435;	Conservative 0;	Mismatches 33;	Indels 0;	Gaps 0;
QY	126	CGCGCCTGCTAGCCCAATGCTCCGATCTTAATGATGAAAGGACCCCTCCCGAGAAAT	185		
DB	490	CCAGCCTGCTAGCCCTTGTTCGATGTTAATGATCAAAAGGACCCCTCTCAGGAAAT	431		
QY	186	CTCACTGACACACCCCTACTATGCCCAATTCAGCGGAGAGCAGTTAGAGCGGTCA	245		
DB	430	CTCANCTGCACACCCCTACCATTGCCCATTTTCAGCAGGAAGCAGGTAGAGCGGTCT	371		
QY	246	GCCAACTCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGCAGGAC	305		
DB	370	GTCAACTCTCCCAAGCAGCAGTGGGTTTCTGTTGAGAGGGGGTACTGAGAGCAGGAC	311		
QY	306	TAGCTGGATTCTCTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAGGTGATGTCATC	365		
DB	310	TAGCTGGATTCTCTAGGCCGACTAAGAAATCCCTAAGCCTAGCTGGGAGGTGACCGGCTC	251		
QY	366	CACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAA	425		
DB	250	CACCTTTAAACACCGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAA	191		
QY	426	AATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCCTGAGAGC	485		
DB	190	AATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCCTGAGAGC	131		
QY	486	ACAGCGGGGAGGACAAAGGATCGGGATATAAACCCAGCAGTTCGAGCGGCAACGCAACC	545		
DB	130	ACAGCAAGAGGAGCAATGATCGGATATAAACCCAGCAGTTCGAGCGGCAACGCGCTACC	71		
QY	546	CCCTTTGGGTCCTCCCTTTGATGGCGCTCTGTTTTCACCTCTATT 593			
DB	70	CTCTTTGGGTCCTCCCTTTGATGGAGCGCTGTTTTCACCTCTATT 23			

RESULT 21

BE732673
 LOCUS
 DEFINITION
 BE732673
 BE732673.1 GI:10146665
 EST.
 KEYWORDS
 SOURCE
 Homo sapiens (human)

BE732673 440 bp mRNA linear EST 15-SEP-2000
 601571305F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3925728 5', mRNA sequence.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 440)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM752 row: j column: 01
High quality sequence stop: 440.

FEATURES
source
1..440
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3925728"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Query Match 62.6%; Score 397.4; DB 2; Length 440;
Best Local Similarity 94.1%; Pred. No. 7 4e-104;
Matches 413; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 179 AGGAAATCTCAATGACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCG 238
DB 1 AGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCGGGAAGCAGTTAGAGCG 60
QY 239 GTCATCAGCAACTCCCAACAGCACTTGGGTTTCTCTGTGAGAGGGGCACTGAGAG 298
DB 61 GTCGTCGGCAACTCCCAACAGCACTTGGGTTTCTCTGTGAGAGGGGCACTGAGAG 120
QY 299 ACAGGACTAGCTGGATTTCCTAGGCCAAACGAAGATCCCTTAAGCCTAGCTGGGAAGGTGA 358
DB 121 ACAGGACTAGCTGGATTTCCTAGGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGA 180
QY 359 CTCGATCCACTCTAAACATGCGGGCTTGAACCTTAGCTACACCCGACCAATCAGAGAGC 418
DB 181 CCACATCCACTTAAACAGCGGGCTTGAACCTTAGCTACACCTGACCAATCAGAGAGC 240
QY 419 TCCTAAATGCTAATTAGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCC 478
DB 241 TCCTAAATGCTAATTAGCAAAAGCAGAGGTAAGAAATAGCAATCATCTATTGCC 300
QY 479 TGAGAGCAGCGGGAGGGAACAGGATCGGGATATAAACCCAGGCATTCGAGCGGCAAC 538
DB 301 TGAGAGCAGCAGGAGGACATGATCGGGATATAAACCCAGTCCTTCGAGCGGCAAC 360
QY 539 GGCAACCCCTTTGGGTCCTCCCTTTGATGCGGCGCTCTGTTTTCACCTCTATTTCACCT 598
DB 361 GGCAACCCCTTTGGGTCCTCCCTTTGATGCGGAGCTCTGTTTTCATGCTATTTCACCT 420
QY 599 CTATTAAATCTTGCACCTG 617
DB 421 CTATTAAATCTTGCACCTG 439

RESULT 22
BI963185/c
LOCUS
DEFINITION i57c10.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5670930 3', mRNA sequence.
ACCESSION BI963185
VERSION BI963185.1 GI:16337590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,
Williams, T., Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 451.

FEATURES
Location/Qualifiers
1..543
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5670930"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

ORIGIN
Query Match 62.4%; Score 396; DB 4; Length 543;
Best Local Similarity 88.3%; Pred. No. 2e-103;
Matches 467; Conservative 0; Mismatches 51; Indels 11; Gaps 3;
QY 74 AAATGGAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGCTG 133
DB 543 AAATGGAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGCTG 484
QY 134 CTAGGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAGGAAATCTCAACTG 193
DB 483 CTAGGCCA-GCTCCACATTAATGACATCAAGGACACCCCTCCCGAGGAAATCTCAACTG 425

```

QY 194 CACAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGTTCATCAGCCAACT 253
Db 424 CATGACACCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTGTGGCCAACT 365
QY 254 CCCCAACAGACACTTGGTGTTCCTGTTGAGGGGGGACTGAGAGACAGACTAGCTGGA 313
Db 364 CCCCAACAGACACTTGGTGTTCCTGTTGAGTGGGGGACTGAGAGACAGACTAGCTGGA 306
QY 314 TTTCTTAGGCAAGGAATCCCTAGCCTAGCTAGCTGGGAAGGTGACTGCATCCACTCTA 373
Db 305 TTTCTTAGGCAACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACTCTA 246
QY 374 AACATGGGCTTGCACTTAGCTACACCCCGACCAAT-----CAGAGAGCTCACTA 424
Db 245 ACACTAGGCTTGCACTTAGCTACACCCCGACCAANTAGGAGTAGTAAGAGAGCTTGCTA 186
QY 425 AAATGCTAATTAGGCAAAATAGAGAGTAAGAAATAGCCAAATCATCTATTGCTGAGAG 484
Db 185 AAATGCTAATTAGGCAAAACAGAGAGTAAGAAATAGCCAGTCATCTATCGCTGACAG 126
QY 485 CACAGCGGAGGACAGAGATCGGGATATAAACCAGGCAATTCAGCGGCAAGGCAAC 544
Db 125 CACAAGGGGGGAGCAATATGATCAGGATATAAATCAAGCTCAGGCAATCAAGCCAGCAATGGCTAC 66
QY 545 CCCCTTTGGGTCCCTCCCTCTGTTGATGGCGCTCTGTTTTCACCTCTATT 593
Db 65 CCACCTTTGGGTCCCTCCCTCTGTTTATGGGAGCTCTGTTTTCACCTCTATT 17

```

RESULT 23

```

BX380176/c
LOCUS BX380176 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSODI044YK06 3-PRIME, mRNA sequence.
ACCESSION BX380176
VERSION BX380176.2 GI:46834414
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

REFERENCE

```

AUTHORS Li W.B., Gruber C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30460243.
Contact: Genoscope

```

```

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r

```

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?b=CSODI044BF03NP1&c=4215.r>.

FEATURES

source

```

1. .966
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI044YK06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notice="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match 62.2%; Score 395.2; DB 5; Length 966;
Best Local Similarity 79.6%; Pred. NO. 3.8e-103;
Matches 465; Conservative 22; Mismatches 84; Indels 13; Gaps 2;

QY 1 CCTGTATCTTTAAACCTCTCTTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60
Db 574 CCTGTATCTTTAAACCTCTCTTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 515
QY 61 CAAATTTGTTCTTTCAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGAGACC 120
Db 514 -----CAAATGGAGCCAGATGCAGTCCARGACTAGGDTCTACCGCGGCC 467
QY 121 CTGACACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTCGAAGCACCCTCCCGAG 180
Db 466 CTGACACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTCGAAGCACCCTCCCGAG 407
QY 181 GAAATCTCAACTGCACAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 406 GAAATCTCAACTGCACAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 347
QY 241 CATCAGCCAACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 346 CGTCGCCCAACCTCCCGACGACCTTGGGTTTCTGTTGAGATGGGGAGCTGAGAGAC 287
QY 301 AGGACTAGCTGGATTTCTAGGCCAAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACT 360
Db 286 MGGACTGGCTGGATTTCTAGGCTGACTAAGAAATCCCTTARGCCTRGCTGGGAAGGTGACC 227
QY 361 GCATCCACCTCTAAACATGGGGCTTCAACTTACCTACACCCGACCAATCAGAGAGCTC 420
Db 226 ACATCCACCTTTAAACMCGGGGCTTGCCACTTGGCTCACCCCTGCCCAATTCRGGGGCTC 167
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGC-CAATCATCTATTGCTCT 479
Db 166 ACTAAATGCTAATTCGGCCCCSCSSSSSVSSGCCCTCGCGCHSYCMTCATTGCTCT 107
QY 480 GAGAGCACGCGGGAGGACAGGATCGGATATAAACCAGGCAATCCAGGCAATCCAGCGGCAACG 539
Db 106 GAGAGCACGCMGAGGAGCAATGATCGGATATAAACCAGGCAATCCAGGCAATCCAGCGGCGCCG 47
QY 540 GCACACCCCTTTGGGTCCCTCCCTTGTATGGCGCTCTGTTT 583
Db 46 GCCCCCCCTTTGGGTCCCTCCCTTGTATGGCGCTCTGTTT 3

RESULT 24
AG058970/c
LOCUS Pan troglodytes DNA, clone: PTB-046A08.R, genomic survey sequence.
DEFINITION AG058970
ACCESSION AG058970.1 GI:16596431
VERSION GSS.
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Pan troglodytes
REFERENCE 1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Shuhiro-chou, Tsukuba, Ibaraki, Japan, 305-8565, Japan
(E-mail: chimpbes@gs.c.riken.go.jp, URL: http://hgp.gs.c.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.

```

```

PRIMERS
Sequencing: M13Rev
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .651
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-04A08.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
ORIGIN
Query Match      62.2%; Score 394.8; DB 9; Length 651;
Best Local Similarity 87.6%; Pred. No. 4.6e-103;
Matches 503; Conservative 0; Mismatches 57; Indels 14; Gaps 6;

QY 20 TTGTTAAGTTTGCTCTTCCCA-GAATCAAAACTGTAAACTACAAATTTGTTCTTCAAATG 78
DB 613 TTGATAGTTTGCTCTTCCAGATCAAGCTGTAAACTACAAATGTTTCTTCAAATG 554

QY 79 GAGCACGATGAGTCCATGAC-TAAGATCCACCGTGGACCCCTGGACCGCCTGCTAG 137
DB 553 GAGCACGATACAGTGCATGTCTTAAGTTCTCCCGTGAACCCCTGGATGGCCAGTAG 494

QY 138 CCATGCTCGATGTTAATGACATTCGAAGGCCCTCCCGAGGAAATCTCAACTGCACA 197
DB 493 CCATGCTCGATGTTAATGACATTCGAAGGCCCTCCCGAGGAAATCTCAACTGCACA 434

QY 198 ACCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGTCTATCAGCCAACTCCCC 257
DB 433 ACCCTGCTACACCCCAATTCAGCAGGAAGCAGTTAGACGAGTCTGTCAGCCAACTCCCC 374

QY 258 AACAGCACTGGGTTTCTTGTGAGAGGGGAGCTGAGACGAGCACTAGCTGGATTTC 317
DB 373 AACAGCACTGGGTTTCTTGTGAGAGTGGGACTGAGACGAGCACTAGCTGGATTTC 314

QY 318 CTAGGCAACGAAGATCCCTAAGCCTAGCT-GGGAAGGTGACTGCATCCACTTAAAC 376
DB 313 CTAGGCTGACTAAGATTCCTAAGCCTAGCTGGGGAAGGTGATTGCACCCACTTAAAC 254

QY 377 ATGGGGCTTGCAACTTAGCTTCAACCCGACCAATCAG-----AGAGCTCACTAAA 427
DB 253 ATGGGGCTTGTAATCAGCTCACACCCCAACAAATCAGGTAGTAAAAAGGGCTCACTAAA 194

QY 428 TGCTAATTAGGC-AAAAATAGGAGTTAAAGAAATAG-CCAATCATCTATTGCTTGAGAGC 485
DB 193 TACAAATTAGCTTAAAGACGAGGAGATTAAGAAATAGTCAANTCATATACCTGAGACC 134

QY 486 ACAGCGGGAGGACAAGGATCGGGATATAAACCCAGGCATTCGAGCGGCAACGGCAACC 545
DB 133 ACAGTGGGAGGACAATGGTCGGGGTATAAACCCAGGCATTCGAGCAGGAGTGGCAACC 74

QY 546 CCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCT 579
DB 73 CGCTTTGGGTCCCTCCCTCAATTGTTATGGGAGCTCT 40

RESULT 25
AG126669/c
LOCUS
DEFINITION Pan troglodytes DNA, clone: PTB-137E19.F, genomic survey sequence.
ACCESSION AG126669
VERSION AG126669.1 GI:16655834
KEYWORDS GSS.
SOURCE
ORGANISM Pan troglodytes (chimpanzee)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1
REFERENCE
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 701)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
Location/Qualifiers
1. .701
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-137E19.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      61.9%; Score 393.2; DB 9; Length 701;
Best Local Similarity 88.1%; Pred. No. 1.4e-102;
Matches 465; Conservative 0; Mismatches 53; Indels 10; Gaps 3;

QY 79 GAGCACGATGGAGTCCATGACTTAAGATCCACCGTGGACCCCTGGACCGCCTGCTAGC 138
DB 651 GACCCGAGATCCAGTCCATGACTTAAAGACTACTGTGAACCCCTGGACTGGCCTGCTAGC 592

QY 139 CCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCT-CCCGAGGAATCTCAACTGCACA 197
DB 591 CCATGCTGCAATGTTAATGACATTTGAAGGCAACCCCTGCCCCGAGGAATCTCAACTGCACA 532

QY 198 ACCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGTCTATCAGCCAACTCCCC 257
DB 531 ACCCTACTATGCCCGATTCAGCAGGAGGCAGTTAAAGTGGTGGTGGCCAACTCCCC 472

QY 258 AACAGCACTTGGGTTTCTTGTGAGAGGGGAGCTAGAGACAGGACTAGCTGGATTTC 317
DB 471 AACAGCACTTGGTTTTCTTGTGAGAGGGGATACTAGAGACAGGACTAGCTGGATTTC 412

QY 318 CTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACTGCATCCACTTAAACA 377
DB 411 CTAGGCTCACTAAGAAATCCCTAAGACTAGCTGGGAGGTGACCATCCACTTGAACA 352

QY 378 TGGGGCTTTGCAACTTAGCTTCACACCCGACCAATC-----AGAGACTCACTAAAATG 429
DB 351 CGGGCTTTGCAACTTAGCTTCACATCCGACCAATCAGGTAATAAGAAAGTAACATAAATG 292

QY 430 CTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTTGAGAGCACAG 489
DB 291 CTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATCATCTAGATCACAG 232

QY 490 CGGAGGAGGACAGGATCGGGATATAAACCCAGGCATTCGAGCGGCAACGGCAACCCCT 549
DB 231 CAGGAGGAGACATGATCGGGATATAAACCCAGGCATTCGAGCGGCAACGACGACTACCTCT 172

QY 550 TTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTTCACCTCTATTTCAC 597
DB 171 TT-GGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCACCTCTATTAAAC 125
```

```

RESULT 26
AG066901/c
LOCUS
DEFINITION
Pan troglodytes DNA, clone: PTB-056L15.F, genomic survey sequence.
ACCESSION
AG066901
VERSION
AG066901.1 GI:16618703
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 609)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teikoku-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpan@psc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Pan troglodytes
Location/Qualifiers
1..609
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-056L15.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 61.9%; Score 393; DB 9; Length 609;
Best Local Similarity 87.7%; Pred. No. 1.5e-102;
Matches 478; Conservative 0; Mismatches 55; Indels 12; Gaps 4;

Qy 47 AAACGTAAACTACAAATTTCTTCAATGGAGCACCGATGGAGTCCATGACTAAGA 106
Db |||
Qy 606 AAGCTGTAGAACTACAAATCATTTCTTCAATGGAGCCCCAAATGCAGTCCACTAAGA 547
Db |||
Qy 107 TCACCGTGGCCCTGGACCGGCTCTAGCCCATCTCC-GATGTTAATGACATTGAA 165
Db |||
Qy 546 TCTACTGCAGACCCCTTGACCGAGCTGTAGCTCATGCTCCAAATGTAATGGCCTTGA 487
Db |||
Qy 166 GGACACC-CCTCCGAGGAAATCTCAACTGCACACACCCCTACTATGCCCAATTCAGCGGG 224
Db |||
Qy 486 AGCCCTTCTCCGAGGAAATCTCACTGACACACCCCTACTACACCCCAATTCAGCAGG 427
Db |||
Qy 225 AAGCAGTTAGACGGTCAATAGCAACCTCCCAACAGCACTGGGTTTCTCTGTTAGA 284
Db |||
Qy 426 AAGCAGTTAGACGAGTATCGGCCAACCTCCCAACAGCACTGGGTTTCTCATGTTAGA 367
Db |||
Qy 285 GGGGGACTGAGACAGGACTAGCTGATTTCTTAGCCCAACGAGATCCCTAAGCCT 344
Db |||
Qy 366 TGGGGGACTTGAACAGGACTAGCTGATTTCTTA-GCTGACTAAGAAATCCCTAAGCCT 308
Db |||
Qy 345 AGCTGGGAAGTGAATGATCCACTCTTAAACATGGGGCTTGCAACTAGCTCACCCG 404
Db |||

```

```

Db 307 ACCTGGGAAGGTGACACACTTCCACCTTTAAACACGGGGCTTGCACCTTAGCTCACACATG 248
Qy 405 ACCAATCA-----GAGAGCTCACTAAAAATGCTAATTAGGCAAAAAATAGAGGTTAA 455
Db |||
Db 247 ACCAATCAGATAGTAGGAGAGCTCACTAAATGCTAATTAGGCAACAACAGAGGTTAA 188
Qy 456 GAAATAGCCAATCATCTATTGCTGTAGAGCACAGCGGGAGGACAGGATCGGATATAA 515
Db |||
Db 187 GAAATAGCCAATCATCTGCTACCTGAAGACACAGTGGGAGGACAAATGATTGGATATAA 128
Qy 516 ACCCAGGCATTCAGCGCCGCAACGCCCTTTGGGTCCCTCCCTTTGTTATGGCG 575
Db |||
Db 127 ACCCAGGCATTTGAGCCAGCAATGGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGGAG 68
Qy 576 CTCTG 580
Db |||
Db 67 CTCCG 63

RESULT 27
AA837267/c
LOCUS
DEFINITION
AA837267 443 bp mRNA linear EST 31-MAR-1998
od26b10.s1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:1369051
similar to contains PTR7.tl PTR7 repetitive element ;, mRNA
sequence.
ACCESSION
AA837267
VERSION
AA837267.1 GI:2912466
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 443)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
AUTHORS
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1375 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham.
FEATURES
Location/Qualifiers
1..443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1369051"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI CGAP GCBI"
/note="Vector: pTR3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo (dT) primer
[5'-TGTTACCAATCTGAGTGGAGCGGCTCATTTTCTTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTR3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN

```

Fri Feb 25 16:26:31 2005

FEATURES		Location/Qualifiers	
source		1..712	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/clones="Plate=3131 Col=8 Row=J"	
		/sex="male"	
		/clone_lib="CIT Approved Human Genomic Sperm Library D"	
		/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"	
ORIGIN			
Query Match		61.0%; Score 387.2; DB 8; Length 712;	
Best Local Similarity		82.2%; Pred. No. 7.5e-101;	
Matches 484; Conservative		0; Mismatches 94; Indels 11; Gaps 3;	
QY	184	ATCTCAACTGCAACACCCCTACTATGCCCCCAATTCAGCGGGAGCAGTTAGAGCGGTCA	243
Db	443	ATCTCAACTGCAACACCCCTACTATGCCCCCAATTCAGCGGGAGCAGTTAGAGCGGTCA	384
QY	244	CAGCCAACTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGACTGAGAGACAGG	303
Db	383	CAGCCAACTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGGACTGAGAGACAGG	324
QY	304	ACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGGAAGTGACTGCA	363
Db	323	ACTAGCTGGATTTCCTAGGCCGAATTAAGAATCCCTAAGCCTAGCTGGGGAAGTGACCGCA	264
QY	364	TCCACTCTTAACATGGGGCTTGCACCTAGCTGACACCCGACCAATCAGAGAGCTCACT	423
Db	263	TCCACTCTTAACATGGGGCTTGCACCTAGCTGACACCCGACCAATCAGAGAGCTCACT	204
QY	424	AAATCTCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGA	483
Db	203	AAATCTCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGA	144
QY	484	GCACAGCGGAGGACAGAGATCGGATATAAACCCAGCAGCATTCGAGCCGGCAACGGCAA	543
Db	143	GCACAGTGGGAGGACAGAGATTGCAATATAAACCCAGCAGCATTCGAGCCGAGC-ANAGCAA	85
QY	544	CCCCCTTTGGGTCCTCCCTTTGATGGCGCTCTGTTTTCACCTCTATTTCACCTCTATT	603
Db	84	CCGCCCTTTGGTCCCTTCCCTTGTATGGAGCTCTGTTTTCACCTCTATTTCACCTCTATT	25
QY	604	AAATCTTGCACTGAAAAAATAA 627	
Db	24	AAATCTTGCACTGAAAAAATAA 1	
RESULT 28		712 bp DNA linear GSS 10-NOV-1999	
AQ892947/c			
LOCUS		HS 3131 B2 E04 MR CIT Approved Human Genomic Sperm Library D Homo	
DEFINITION		sapiens_genomic clone Plate=3131 Col=8 Row=J, genomic survey	
sequence.			
AQ892947		GI:6349137	
VERSION		GSS.	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS		1 (bases 1 to 712)	
TITLE		Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,	
JOURNAL		Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and	
MEDLINE		Hood,L.	
COMMENT		Sequence-tagged connectors: A sequence approach to mapping and	
		scanning the human genome	
		Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)	
		99380589	
		10449764	
		Contact: Mahairas GG, Wallace JC, Hood L	
		High Throughput Sequencing Center	
		University of Washington	
		401 Queen Anne Avenue North, Seattle, WA 98109, USA	
		Tel: (206) 616-3618	
		Fax: (206) 616-3887	
		Email: jwallace@u.washington.edu	
		Clones may be purchased from Research Genetics (info@resgen.com).	
		BAC end Web Server: http://www.htsc.washington.edu	
		Plate: 3131 row: J column: 8	
		Seq primer: M13 Reverse	
		Class: BAC ends	
		High quality sequence stop: 712.	
AL128526		436 bp mRNA linear EST 27-OCT-1998	
qc61h10.x1 Soares_placenta_8to9weeks_2NBHP8to9W Homo sapiens cDNA			
clone IMAGE:1714147 3', similar to contains PTR7.bl PTR5 repetitive			
element ;, mRNA sequence.			
AI128526		GI:3597040	
AI128526.1		EST.	
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
1 (bases 1 to 436)			
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			

JOURNAL
COMMENT

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 720 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 428.

FEATURES
source
1. .436
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:171417"
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta 809weeks 2NbhP8to9W"
/notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGGATTTTTTTTTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bernaldo."

ORIGIN

Query Match 60.9%; Score 386.8; DB 1; Length 436;
Best Local Similarity 93.7%; Pred. No. 8.8e-101; Indels 0; Gaps 0;
Matches 403; Conservative 0; Mismatches 27;

QY 206 TATCCCCCAATTTCAGCGGAGCAGTGTAGAGCGGTCTATCGCCCAACCTCCCCAACAGCAC 265
Db 436 TACGCCCAATTTCAGCAGGAAGCAGTTAGAGCGGTCTCGGCCCAACCTCCCCAACAGCAC 377

QY 266 TTGGGTTTTCTGTGTAGAGGGGGGACTGAGACAGGACTAGCTGTGATTTCTTAGGCCA 325
Db 376 TTAGGTTTTCTGTGTAGATGGGGGACTGAAAGACAGGACTAGCTGTGATTTCTTAGGCTG 317

QY 326 ACGAGAGATCCCTAAGCTAGCTGGAGGAGTGTGATCCACCTCTAAACATGGGGCTT 385
Db 316 ACTAAGAGATCCCTAAGCTAGCTGGAGGAGTGTGATCCACCTCTAAACATGGGGCTT 257

QY 386 GCAACTTAGCTCACACCCGCAATCATCAGAGAGCTCAGTAAATGCTTAATAGGCAAAAT 445
Db 256 GCAACTTAGCTCACACCTGACCATCATCAGAGAGCTCAGTAAATGCTTAATAGGCAAAAG 197

QY 446 AGGAGGTAAGAAATAGCAATCATCTATTGCTTGTAGAGCAGCGGGAGGAGCAAGAT 505
Db 196 AGGAGGTAAGAAATAGCAATCATCTATTGCTTGTAGAGCAGCGGGAGGAGCAATGAT 137

QY 506 CGGATATTAACCCAGGATTCGAGCGGCAACCGCAACCCCTTTGGGTCCCTCCCTT 565
Db 136 CGGATATTAACCCAGGATTCGAGCGGCAACCGCAACCCCTTTGGGTCCCTCCCTT 77

QY 566 TGTATGGCGCTGTGTTTCACTCTATTTCATCTATTAAATCTTTGCAACTGAAAAAAA 625
Db 76 TGTATGGAGCTGTGTTTTCATGCTATTTCATCTATTAAATCTTTGCAACTGAAAAAAA 17

QY 626 AAAAAAAA 635
Db 16 AAAAAAAA 7

RESULT 30
AI393478/c
LOCUS
DEFINITION
IMAGE:2111766 3', mRNA sequence.
ACCESSION
AI393478

AI393478.1 GI:4223025
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 446)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1061 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 445.

FEATURES
source
1. .446
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2111766"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
1.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bernaldo."

ORIGIN

Query Match 60.7%; Score 385.2; DB 1; Length 446;
Best Local Similarity 92.5%; Pred. No. 2.6e-100; Indels 10; Gaps 1;
Matches 419; Conservative 0; Mismatches 24;

QY 183 AATCTCACTGACCAACCCCTACTATGCCCAATTCAGCGGAGCAGTGTAGAGCGGTCA 242
Db 446 AATTTCACTGTGCACCCCTANTATGCCCAATTCAGCAGGAAGCAGTTACAGCGGTCA 387

QY 243 TCAGCCCAACCTCCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGAGCTGAGAGACAG 302
Db 386 TCAGCCCAACCTCCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGAGGGGCACTGAGAGACAA 327

QY 303 GACTAGCTGATTTCTTAGGCCAACGAGAAATCCCTAAGCCTAGCTGGAGGAGTACTGC 362
Db 326 GACTAGCTGATTTCTTAGGCCGAGTAAAGAAATCCCTAAGCGTAGCTGGAGGAGTACTGC 267

QY 363 ATCCACCTCTAAACATGGGCTTGCACCTTAGCTCACACCCGCAATTCAGAGAGCTCAC 422
Db 266 ATCCACCTTTAAACATGGGCTTGCACCTTAGCTCACACCCGCAATTCAGAGAGCTCAC 207

QY 423 TAAATGCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTGTAG 482
Db 206 TAAATGCTTAATAGGCAAAACAGGAGGTAAAGAAATAGCAATCATCTATTGCTGTAG 147

QY 483 AGCACAGGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAACGCA 542
Db 146 AGCACAGGAGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAATTCGCA 87

QY 543 ACCCCCTTTGGTCCCTCTTTGATGGCGCTCTGTTTTCATCTATTTCATCTCTAT 602
Db 86 ACCCCCTTTGGTCCCTCTTTGATGGAGCTCTGT-----TTTCACTCTAT 37

QY 603 TAAATCTTGCACTGAAAAA 635

Db 480 CACATGCTCGATGTTAATGACATCGAAGGACCCGCTCTCGAGGAAATCTCAACTGCACA 421
 QY 198 ACCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGGGTCATCAGCCACCTCCCC 257
 Db 420 ACCCTACTACACCCCAATTCAGCGGGAAGCAGTTAGAGGGTCATCAGCCACCTCCCC 361
 QY 258 AACAGCAGTCTGGGTTTCTGTTGAGAGGGGGGACTCGAGAGCAGGACTAGCTGGATTTC 317
 Db 360 AGCAGCAGTCTGGGTTTCTGTTGAGAGGGGGGACTCGAGAGCAGGACTAGCTGGATTTC 301
 QY 318 CTAGGCAACGAAGAATCCCTAAGCCCTAGCTAGCTGGGAAGGTGACTGCATCCACCTCTAAACA 377
 Db 300 CTAGGCTGACTAAGAATCCCTAAGCCCTAGCTAGCTGGGAAGGTGACTGCATCCCTTTAAACA 241
 QY 378 TGGGGCTTCAACTAGCTACACCCGACCAATC-----AGAGAGCTCACTAAAT 428
 Db 240 CGGGCTTCAACTAGCTACACCCGACCAATCAGATAGTAAGAGAGCTCACTAAAT 181
 QY 429 GCTAATTAGGCAAAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGCACA 488
 Db 180 GCTAATTAGGCAAAAAATAGGAGGTAGAGAAATAGCCCAATCATCTATCGCTGAGAGCACA 121
 QY 489 GCGGAGGGAACAAGGATCGGATATAAACCCAGGATTCGAGCGGCAACGGCAACCCCC 548
 Db 120 GCAGGAGGGAACAATGATCCGATATAAACCAAGCATTCGAGCGGCAATGGCTACCCCTC 61
 QY 549 TTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATT 593
 Db 60 TTTGTGTCCTCCCTTTGATGGAGCTCTATTTTCACTCTATT 16

RESULT 33
 BX481837/c
 LOCUS 664 bp mRNA linear EST 04-SEP-2003
 DEFINITION DKFZp686M05227 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 DKFZp686M05227 5', mRNA sequence.
 BX481837
 BX481837.1 GI:31941162
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 664)
 Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
 Fobo, G., Han, M. and Wiemann, S.
 EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
 Unpublished (2003)
 Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No sl sequence available.
 This clone (DKFZp686M05227) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1..664
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686M05227"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hlcc3)"
 /note="vector: pTriplex2; Site_1: SfIIA; Site_2: SfiIB;
 cDNA-collection"

ORIGIN

Query Match 59.8%; Score 379.8; DB 5; Length 664;
 Best Local Similarity 84.3%; Pred. No. 1e-98;
 Matches 514; Conservative 0; Mismatches 78; Indels 18; Gaps 7;
 QY 40 AGAATCAAACTGTAAACTACAAATTTCTTTCAATGGAGCCAGATGGAGTCCATG 99
 Db 664 AGAATCAAACTGTAAANACACATCGTTCTTCAATGGAGCCAGATGGAGTCCATG 605
 QY 100 ACTAAGATCCACCTGGAGCCCTCGACCGGCTGCTAGCCCATGCTCCGATGTTAATGAC 159
 Db 604 ACTAAGATCTACCGCGGATCCCTGGACGAGCTGCTAGCCCATGCTCCGATGTTAATGAC 545
 QY 160 ATTAGAGGACCCCTCCCGAGGAATCTCAACTGCACAACTCTACTATGCCCAATTC 219
 Db 544 ATTAGAGGACCTCTCCCGAGGAATCTCAAAATGCACAACTCTACTATGCCCAATTC 485
 QY 220 GCGGAGGAGCTAGAGCGGTCTCAGCCAACTCCCAACAGCAGCTTCGGTTCCTGT 279
 Db 484 GCGGAGGAGCTAGAGCGGTCTCAGCCAACTCCCAACAGCAGCTTCGGTTCCTGT 425
 QY 280 TGAGAGGGGGGACTCGAGAGCAGGACTAGCTGGATTCTTAGCCCAACGAAGATCCCTA 339
 Db 424 TGAGAGGGGGGACTCGAGAGCAGGACTAGCTGGATTCTTAGCCCAACGAAGATCCCTA 365
 QY 340 AGCTTAGCT-GGGAGGTTGATCTGATCTCACTTAAACATGGGGC--TTGCACTTAGCT 396
 Db 364 AGCTTAGTGGGGAAGGTGACTGACCTTAAACACAGGACTTTTGTAACTCAGCT 305
 QY 397 CACACCCGACCAATC-----AGAGAGCTCACTATAATGCTAATAGGC-AAAAATA 446
 Db 304 CACATCTGCCCAATCAGGTAGTAAGAGAGAGCTATTAAATATCAATATTAGCTTAAACA 245
 QY 447 GGAGGTAAAGAAATAGCCCAATCATCTATTGCTCTGAGAGCAGCAGCGGAGGAGCAAGATC 506
 Db 244 GGAGGTAAAGAAATA---ATCATCTCTGCTCTGAGAGCAGCAGAGGAGGCAATGATT 188
 QY 507 GGGATATAAA-CCAGGCAATTCAGCGGCA-ACGGCAACCCCTTTGGTGGTCCCTCCCT 564
 Db 187 GGGATATAAACCCAGGCAATTCAGCTGGAGTGGGCAACCCCTTTGGTGGTCCCTCCCA 128
 QY 565 TTGTATGGGGCTCTGTTTCACTCTATTTCATCTATTATAATCTTGCACACTGAAAAA 624
 Db 127 TTGTATGGGAGCTCTGTTTCACTCTATTTCATCTATTTCACACTGCAAAAAA 68
 QY 625 AAAAAA 634
 Db 67 AAAAAA 58

RESULT 34
 AU158595/c
 LOCUS 556 bp mRNA linear EST 05-AUG-2002
 DEFINITION AU158595 PLACE3 Homo sapiens cDNA clone PLACE3000382 3', mRNA
 sequence.
 AU158595
 AU158595.1 GI:11020116
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 556)
 Ota, T., Suzuki, Y., Saito, K., Iishi, S., Yamamoto, J., Sugiyama, T.,
 Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y. and Isogai, T.
 HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Iishi, S.,
 Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
 Masuho, Y., Isogai, T.)
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975

Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES
 source Location/Qualifiers
 1..556
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE300382"
 /tissue_type="Placenta"
 /clone_lib="PLACE3"
 /note="Vector: pME18SFL3"

ORIGIN
 Query Match 59.1%; Score 375.2; DB 1; Length 556;
 Best Local Similarity 84.1%; Pred. No. 2.1e-97;
 Matches 470; Conservative 0; Mismatches 50; Indels 39; Gaps 3;

48 AACCTGTAATAACTCAAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGAT 107
 |||||
 540 AACTNGTAATACTTCAATGTTCTTCAA- GGAGTCCCGATNCAGTCCATGATTAAAGC 482
 |||||
 108 CCACCGTGGACCCCTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTGAAGG 167
 |||||
 481 CTACNGTGNACCTTGGACTGGGCTGTTAGCCCATGTTCTGATGTTAATGACATCAAAGG 422
 |||||
 168 CACCCTCCGAGGAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGGAAG 227
 |||||
 421 CATCCTCCGAGGAAATCTCAACTGCACACCCCTANTATGCCCAATTCAGCAGAAG 362
 |||||
 228 CAGTTAGAGCGGTTCATCAGCCAACTCTCCCAACAGACACTTGGGTTTTCTTGTGAGAGG 287
 |||||
 361 CAGTTAGAGCGGTTCATCGGCCAACCTCCCAACAGACACTTGGGTTTTCTTGTGANAGG 302
 |||||
 288 GGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAAGAGAAATCCCTAAGCGTAGC 347
 |||||
 301 GGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGAGTAAAGAAATCCCTAAGCGTAGC 242
 |||||
 348 TGGGAAGTGACTGCATCCACCTCTAAACATGGGGCTTCGAACTTAGCTCAACCCGACC 407
 |||||
 241 TGGGAAGTGACACATCCACCTTTAAACACAGGGCTTGCACCTTAGCTCAACCCGACC 182
 |||||
 408 AATC-----AGAGAGCTCACTAAATGCTAAATAGGCGAAAAATAGGAGTTAAGAA 458
 |||||
 181 AATCAGGTAGCAAGAGAGCTCGTTAAATGATAATAGGCGAAAAACAGGAGTTAAGAA 122
 |||||
 459 ATAGCAATCATTTATTCCTGAGAGACACAGCGGAGGCAAGGATCGGATATAAACC 518
 |||||
 121 ATAGCAATCATTTATTCGCTGAGAGCACAG----- 91
 |||||
 519 CAGGCAATTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCTTTGTATGGCGCTC 578
 |||||
 90 CAGGCAATTCGAGCGGCAATGGCAACCTCTTTGGGTCTCTCTCTTTGTATGGCGCTC 31
 |||||
 579 TGTTTTCACTCTATTTCAC 597
 |||||
 30 TGTTTTCACTCTATTAAAC 12
 |||||

RESULT 35
 CB231128/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CB231128 865 bp mRNA linear EST 10-FEB-2003
 AGENCOURT 11468603 NICHD Rh.Ov1 Macaca mulatta cDNA clone
 IMAGE:688454 5', mRNA sequence.
 CB231128
 CB231128.1 GI:28282706
 EST.
 Macaca mulatta (rhesus monkey)
 Macaca mulatta
 Euteleostomi;

```
RESULT 36
AI074704/c
LOCUS
DEFINITION
AI074704 470 bp mRNA linear EST 06-AUG-1998
ox83d05.s1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1662921
3' similar to contains PTR7.t2 PTR7 repetitive element ;, mRNA
sequence.
ACCESSION
AI074704.1 GI:3401348
VERSION
AI074704.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 470)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 447.
FEATURES
Location/Qualifiers
1..470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1662921"
/tissue_type="pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NhMpu S1"
/notes="Organ: mixed (see below); Vector: p7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not 1;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NhMpu, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
ORIGIN
Query Match 57.6%; Score 365.8; DB 1; Length 470;
Best Local Similarity 90.8%; Pred. No. 1.1e-94;
Matches 415; Conservative 0; Mismatches 32; Indels 10; Gaps 2;
Qy 146 CCGATGTTAATGACATTAAGGACCCCTCCCGAGGAATCTCACTGCACACCCCTAC 205
Db 470 CCGATGTTGATGACATCAAGGACCCCTCTCTGAGGAATCTCACTGCACACCCCTAC 411
Qy 206 TATGCCCAATTCAGCGGAAGAGCTAGAGCGGTATCATGCCAACCTCCCAACAGCAC 265
Db 410 TATGCCCAATTCAGCGGAAGAGCTAGAGCGGTATCATGCCAACCTCCCAACAGCAC 351
Qy 266 TTGGGTTTCTGTTGAGGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCA 325
Db 350 TTGGGTTTCTGTTGAGGTGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTG 291
Qy 326 ACGAAGATCCCTAAGCCCTAGCTGGGAGGTGACTGCATCCACCTCTAAACATGGGGCTT 385
Db 290 ACTAAGATCCCTAAGCCCTAGCTGGGAGGTGACGGCATCCACCTTTAAACACGGGCTT 231
Qy 386 GCAACTTAGCTCACACCCGACCAATC-----AGAGACTCACTAAATGCTAATTA 436
Db 230 GCAACTTAGCTCACACCCGACCAATCAGGGAGTAAAGACAGCTCACTAAATGCTAATTA 171
```

```
Qy 437 GCGAAATAATAGAGGTAAAGAAATAGCCAAATCATCTATTGCTTGAGAGCAGCGGAGG 496
Db 170 GGCATAAACAGGAGGTAAAGAAATAGCCAATCATCTATCGCTGAGAGCAC-GAGGAGG 112
Qy 497 GACAGGATCGGATATAAACCCAGGCATTCCAGCGCGCAACGGCAACCCCTTTGGGTC 556
Db 111 GACAATGATCAGATATAAACCCAGGCATTCAAGCGCGCAGTGGCTACCGCTTTGGGTC 52
Qy 557 CCTCTCCTTTGTATGGCGCTCTCTGTTTTCACCTTATT 593
Db 51 CCTCTGTTTGTATGGGAGCTCTGTTTTCACCTTATT 15
RESULT 37
AG076758/c
LOCUS
DEFINITION
AG076758 679 bp DNA linear GSS 03-NOV-2001
Pan troglodytes DNA, clone: PTB-071A23.R, genomic survey sequence.
ACCESSION
AG076758.1 GI:16628560
VERSION
AG076758.1
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
2 (bases 1 to 679)
Unpublished
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbos@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
FEATURES
Location/Qualifiers
1..679
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-071A23.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 57.4%; Score 364.8; DB 9; Length 679;
Best Local Similarity 90.3%; Pred. No. 2.3e-94;
Matches 390; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 CCTGTATCTTTAACTCCCTTGTAGTTGTCTCTTCCAGAACTCAAACTGTAACCTA 60
Db 475 CCTGTATATTCACCTCTTGTAGTTGTCTCTTCCAGAACTTAAGGCTGTAACCTA 416
Qy 61 CAAATTTCTTCAAATGGAGCACCAAGATGGAGTCCATGACTAAGATCCACCTGGACCC 120
Db 415 CAAATGTTCTTCAAATGAAGCCCAAGATGCAGTCCATGGCTAAATCTACCATGTACCC 356
Qy 121 CTGACCGGGCTGTAGCCCATGCTCGATGTTTAATGACATTTGAGCGACCCCTCCGAG 180
Db 355 CTGGCCAGGCTGTAGCCCTTGTCTCGATGTTTAATGACATCAAAAGCACCCCTCTGAG 296
```

181 GAAATCTCAACTGCAACACCCCTACTATGCCCCCAATTGAGGGGAGCAGTTAGAGCGGT 240
 295 GAAATCTCAACTGCAACACCCCTACTATGCCCCCAATTGAGGGGAGCAGTTAGAGCGGT 236
 241 CATCAGCCAACTCCCAACAGCAGCTTGGGTTTCTGTGAGAGGGGGGAGTCTGAGAGAC 300
 235 CGTCAGTCAACTCCCAACAGCAGCTTGGGTTTCTGTGAGAGGGGGTCTGAGAGAC 176
 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 175 AGGACTAGCTGGATTTCTTAGGCCCAACGAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACC 116
 361 GCATCACCCTTAACATGGGGCTTGGCAACTTAGCTTACACCCGACCAATCAGAGAGCTC 420
 115 GCGTCCACCTTTAAACACGGGGCTTGGCAACTTAGCTTACACCCGACCAATCAGAGAGCTC 56
 421 ACTAAATGCTA 432
 55 TCGAGCATGCTA 44

RESULT 38
 A1128496/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 415)
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert length: 1184 Std Error: 0.00
 Seq primer: -40m13 fwd. Et from Amersham
 High quality sequence stop: 413.

FEATURES
 source
 1. .415
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1714118"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9w"
 /notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGCGCGGATTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo."

ORIGIN
 Query Match 57.4%; Score 364.2; DB 1; Length 415;
 Best Local Similarity 93.7%; Pred. No. 3.1e-94;
 Matches 389; Conservative 0; Mismatches 25; Indels 1; Gaps 1;
 203 TACTATGCCCAATTACGCGGAGCAGTTAGAGCGGTATCAGCCAACTCCCAACAG 262

415 TACTAGCCCCCAATTGAGGAGGAGCAGTTAGAGCGGTCTCGGCCAA-CTCCNCANCAG 357
 263 CACTTGGGTTTCTGTGAGAGGGGGGAGTCTGAGAGCAGAGCTAGCTGAGTTTCCTAGG 322
 356 CACTTAGGTTTCTGTGAGAGTGGGGGAGTCTGAGAGCAGAGCTAGCTGAGTTTCCTAGG 297
 323 CCAACGAAGAATCCCTAAGCCTAGCTGGGAGGTGACTGCATCCACTCTAAACATGGGG 382
 296 CTGACTAAGAATCCCTAAGCCTAGCTGGGAGGTGACCATCCACCTTTAAACACGGGG 237
 383 CTTGCAACTTAGCTCACAACCCGACCAATCAGAGAGCTCCTAAATCTTAATAGGCAAA 442
 236 CTTGCAACTTAGCTCACAACCTGACCAATCAGAGAGCTCCTAAATCTTAATAGGCAAA 177
 443 AATAGGAGGTAAAGAAATAGCCCAATCTATTTCCTGAGAGCAGAGCGGAGGAGCAAG 502
 176 AACAGGAGGTAAAGAAATAGCCCAATCTATTTCCTGAGAGCAGAGCGGAGGAGCAAT 117
 503 GATCGGATATAAACCCAGGCAATTCGAGCGCGCAACCGGCAACCCCTTTGGGTCCCTCC 562
 116 GATCGGATATAAACCCAGGCTTCCTGAGCGCGCAACCGGCAACCCCTTTGGGTCCCTCC 57
 563 CTTTGTATGGCGCTCTGTTTCACTCTATTTCATCTATTAAATCTTGCAACTG 617
 56 CTTTGTATGGAGCTCTGTTTCACTCTATTTCATCTATTAAATCTTGCAACTG 2

RESULT 39
 AG030228/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Pan troglodytes (chimpanzee)
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Tokoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB
 Unpublished
 2 (bases 1 to 777)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Tokoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimpes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS
 Sequencing: M13Rev
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.
 Location/Qualifiers
 1. .777
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-002120.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
 source

ORIGIN

Query Match 56.8%; Score 360.6; DB 9; Length 777;
Best Local Similarity 86.2%; Pred. No. 3.9e-93;
Matches 399; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
Db 515 CCCTATATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAAGCTGCAAACTG 456

QY 61 CAAATTTGTTCTCAATGAGCAGCAGATGAGTTCATGATCAAGTCAAGTCCACGTCGACCC 120
Db 455 CAAATCATCTTCAATGAGCAGCAGATGAGTTCATGATCAAGTTCCTGATGACCC 396

QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTAAATGACATTAAGGACACCCCTCCGAG 180
Db 395 CTGACCGGCTCTAGCCCATGCTCCGATGTTAAATGACATTAAGGACACCCCTCCGAG 336

QY 181 GAAATCTCACTGCACACACCCCTACTATGCCCCCAATTCAGCGGGAAGCATTTAGAGCGGT 240
Db 335 GAAATCTCAATTCGACACCCCGTAAATAGCGCCCAATTCAGCAAGAGCAGTTAGAGCAGT 276

QY 241 CATCAGCAACCTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 275 CGTGGGCAACCTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 216

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAGGCTAGCTGGGAAAGTGACT 360
Db 215 AAGATTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAGGCTAGCTGGGAAAGTGACT 156

QY 361 GCATCCACCTCTAAACATGGGGCTGCACTTACTAGCTCACACCCGACCAATCAGAGAGTTC 420
Db 155 GCATCCACCTTTAAACATGGGGCTGCACTTACTAGCTCACACCCGACCAATCAGAGAGTTC 96

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGC 463
Db 95 TCGAGCATGCTAGCGGGCGCCCTATAGTAAAGTCGTATAACAGC 53

RESULT 40
AI570707/c
LOCUS
DEFINITION
tm7909.x1 NCI_CGAP Brn25 Homo sapiens cDNA clone IMAGE:2164384 3'
Similar to contains PTR7.t1 PTR repetitive element i, mRNA
sequence.

ACCESSION
AI570707.1 GI:4534081
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1. (bases 1 to 425)
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTCAP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1440 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 414.
Location/Qualifiers
1. .425
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2164384"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn25"
/note="Organ: brain; Vector: pTT3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATAGGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pTT3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 56.7%; Score 359.8; DB 1; Length 425;
Best Local Similarity 93.3%; Pred. No. 5.9e-93;
Matches 376; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 191 CTGCACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTATCAGCCAA 250
Db 425 CTGCACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTATCAGCCAA 366

QY 251 CTTCCCAACAGCAGCATTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGACAGGACTAGCT 310
Db 365 CTTCCCAACAGCAGCATTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGACAGGACTAGCT 306

QY 311 GGATTTCTAGGCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACTGCAATCCACCT 370
Db 305 GGATTTCTAGGCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACTGCAATCCACCT 246

QY 371 CTAACATGGGCTTGCACCTTAGCTCACACCGCAATCAGAGAGCTCAGTAAATATGC 430
Db 245 TTAACACGCGGCTTGCAACTTAGCTCACACCGCAATCAGAGAGCTCAGTAAATATGC 186

QY 431 TAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCTGAGAGCAGC 490
Db 185 TAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCTGAGAGCAGC 126

QY 491 GGAGGGAACAAGATCGGATATATAACCCAGGATTCGAGCGCGCAACGCAACCCCTTT 550
Db 125 AAGAGGGAACAAGATCGGATATATAACCCAGGATTCGAGCGCGCAACGCAACCCCTTT 66

QY 551 TGGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATT 593
Db 65 TGGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATT 23

Search completed: February 21, 2005, 14:51:02
Job time : 1996.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:32:54 ; Search time 6356.91 Seconds
(without alignments)
11288.850 Million cell updates/sec

Title: US-09-319-156B-9
Perfect score: 1481
Sequence: 1 agggccctccctatcatcac.....gtcttacaatggaacccca 1481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	1481	6	BD136197
2	1481	100.0	1481	6	AR344387
3	1481	100.0	1481	6	AX001027
4	1481	100.0	1932	14	AF127228
5	1468.2	99.1	162579	9	AL390039
6	1466.6	99.0	2055	6	CQ719381
7	1461.8	98.7	1629	6	BD252040
8	1461.8	98.7	1629	6	AX036992
9	1461.8	98.7	1629	14	AF331500
10	1461.8	98.7	172918	2	AC023366
11	1389.8	93.8	2074	6	AX478550
12	1341.8	90.6	10230	9	AY101589
13	1341.8	90.6	184675	9	AC145964
14	1340.2	90.5	10229	9	AY101586
15	1340.2	90.5	10229	9	AY101587
16	1340.2	90.5	10230	9	AY101588
17	1338.6	90.4	10246	9	AY101593
18	1338.6	90.4	10248	9	AY101592
19	1335.4	90.2	2694	9	AF520478

20	1335.4	90.2	2694	9	AF520487S2	AF520488 Homo sapi
21	1335.4	90.2	2694	9	AF520489S2	AF520490 Homo sapi
22	1335.4	90.2	2694	9	AF520505S2	AF520506 Homo sapi
23	1335.4	90.2	2694	9	AF520517S2	AF520518 Homo sapi
24	1335.4	90.2	2694	9	AF520519S2	AF520520 Homo sapi
25	1335.4	90.2	2694	9	AF520521S2	AF520522 Homo sapi
26	1335.4	90.2	2694	9	AF520529S2	AF520530 Homo sapi
27	1335.4	90.2	2694	9	AF520561S2	AF520562 Homo sapi
28	1335.4	90.2	2930	6	AX355872	AF5205872 Sequence
29	1335.4	90.2	2930	9	AF208161	AF208161 Homo sapi
30	1335.4	90.2	2946	6	AR177269	AR177269 Sequence
31	1335.4	90.2	2946	6	BD232464	BD232464 Secreted
32	1335.4	90.2	2946	6	BD086133	BD086133 Secreted
33	1333.8	90.1	1624	9	AF506835	AF506835 Homo sapi
34	1333.8	90.1	1860	9	AF513360	AF513360 Homo sapi
35	1333.8	90.1	2055	6	BD221827	BD221827 Nucleic s
36	1333.8	90.1	2055	6	AX007999	AX007999 Sequence
37	1333.8	90.1	2599	6	BD221806	BD221806 Nucleic s
38	1333.8	90.1	2599	6	AX007978	AX007978 Sequence
39	1333.8	90.1	2694	9	AF520481S2	AF520482 Homo sapi
40	1333.8	90.1	2694	9	AF520483S2	AF520484 Homo sapi
41	1333.8	90.1	2694	9	AF520485S2	AF520486 Homo sapi
42	1333.8	90.1	2694	9	AF520491S2	AF520492 Homo sapi
43	1333.8	90.1	2694	9	AF520493S2	AF520494 Homo sapi
44	1333.8	90.1	2694	9	AF520495S2	AF520496 Homo sapi
45	1333.8	90.1	2694	9	AF520497S2	AF520498 Homo sapi

ALIGNMENTS

RESULT 1	BD136197	1481 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD136197	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.			
DEFINITION	BD136197	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.			
ACCESSION	BD136197.1	GI:23231142			
VERSION	BD136197.1	GI:23231142			
KEYWORDS	JP 2002509437-A/7.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1481)				
AUTHORS	Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F., Perron,H. and Mandrand,B.				
TITLE	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses				
JOURNAL	Patent: JP 2002509437-A 7 26-MAR-2002;				
COMMENT	BIO MERIEUX				
	OS Unidentified				
	PN JP 2002509437-A/7				
	PD 26-MAR-2002				
	PF 07-JUL-1998 JP 1999508255				
	PR 07-JUL-1997 FR 97/08816				
	PI GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC PI BEDIN,				
	PI MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON,				
	PI BERNARD MANDRAND				
	PC C12N15/48,C12Q1/70,C07K14/15,A61K31/70				
	CC Strandedness: Single;				
	CC Topology: Linear;				
	CC Retroviral nucleic material and nucleotide fragments, in CC particular,				
	CC associated with multiple sclerosis and/or rheumatoid arthritis,				
	CC for				
	CC diagnostic, prophylactic and therapeutic uses FH				
	CC Location/Qualifiers				
	FT source	1..1481			
	FT	/organism='Unidentified'.			
FEATURES		Location/Qualifiers			

```

source      1. .1481
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"

ORIGIN
Query Match      100.0%; Score 1481; DB 6; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCCTCCCTTATCATACTTTTCTTTTACTGTCTCTTACCCTTTTCGTCTCACT 60
    |||||
Db 1 ATGCCCTCCCTTATCATACTTTTCTTTTACTGTCTCTTACCCTTTTCGTCTCACT 60

Qy 61 GCACCCCTCCATGCTGCTGTGAACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
    |||||
Db 61 GCACCCCTCCATGCTGCTGTGAACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120

Qy 121 ACGGGCTTCTTGAAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
    |||||
Db 121 ACGGGCTTCTTGAAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180

Qy 181 ACCTTCACTGCCACAACCATATATGCCCGGCAACTGTCTATACTTGCACCTTTTGCATG 240
    |||||
Db 181 ACCTTCACTGCCACACCATATATGCCCGGCAACTGTCTATACTTGCACCTTTTGCATG 240

Qy 241 CATGCAATACTCATATTATGACAGGGAATAATGATTAACTCTAGTTGTCTTGGAGGACTT 300
    |||||
Db 241 CATGCAATACTCATATTATGACAGGGAATAATGATTAACTCTAGTTGTCTTGGAGGACTT 300

Qy 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCAAGTAGTCTGATGGGGTGGGAATT 360
    |||||
Db 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCAAGTAGTCTGATGGGGTGGGAATT 360

Qy 361 CAAGGTGAGGACAGAGAAAAACAAGTAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420
    |||||
Db 361 CAAGGTGAGGACAGAGAAAAACAAGTAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420

Qy 421 AGCACCCCTAGCCCCCTACAAAGGACTAGTTTCTCTCAAACTACATGAAACCTCCGTACC 480
    |||||
Db 421 AGCACCCCTAGCCCCCTACAAAGGACTAGTTTCTCTCAAACTACATGAAACCTCCGTACC 480

Qy 481 CATACTCGCTGGTGAGGCTATTTAATACGACCCCTCACTCGGTCCATGAGTCTCAGCC 540
    |||||
Db 481 CATACTCGCTGGTGAGGCTATTTAATACGACCCCTCACTCGGTCCATGAGTCTCAGCC 540

Qy 541 CAAPACCTCTACTAATCTGTGATGTGCCTCCCCCTGCACCTTCAAGCCATACATTTCAATC 600
    |||||
Db 541 CAAPACCTCTACTAATCTGTGATGTGCCTCCCCCTGCACCTTCAAGCCATACATTTCAATC 600

Qy 601 CCGTTTCTTGAAACAATCTTGAAACAATCTTGAAACAATCTTGAAACAATCTTGAAACA 660
    |||||
Db 601 CCGTTTCTTGAAACAATCTTGAAACAATCTTGAAACAATCTTGAAACAATCTTGAAACA 660

Qy 661 GGACCTCTGTGTTTCCAACTCTGAAATAACCCATACCTCAAACTCACTGTGTGTAATAATT 720
    |||||
Db 661 GGACCTCTGTGTTTCCAACTCTGAAATAACCCATACCTCAAACTCACTGTGTGTAATAATT 720

Qy 721 AGCAATACTATAGACACAACACGACTCCCAATGCATCAGGTGGGTAAACCTCCCAACGA 780
    |||||
Db 721 AGCAATACTATAGACACAACACGACTCCCAATGCATCAGGTGGGTAAACCTCCCAACGA 780

Qy 781 ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCCTATCATTTGTTG 840
    |||||
Db 781 ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCCTATCATTTGTTG 840

Qy 841 AATGGCTCTTCAGAATCTATGTGCTTCTCTCTCATTTCTTAGTGGCCCTATGACCATCTAC 900
    |||||
Db 841 AATGGCTCTTCAGAATCTATGTGCTTCTCTCTCATTTCTTAGTGGCCCTATGACCATCTAC 900

Qy 901 ACTGAAACAAGATTTTATACAAATATGTGTAACCTTAAGCCCCCAACAAGAGATGCCATT 960
    |||||
Db 901 ACTGAAACAAGATTTTATACAAATATGTGTAACCTTAAGCCCCCAACAAGAGATGCCATT 960

```

Qy	961	CTTCCCTTTTGGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTTGGCATGTATC	1020
Db	961	CTTCCCTTTTGGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTTGGCATGTATC	1020
Qy	1021	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATCGTGACATGGAAACAG	1080
Db	1021	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATCGTGACATGGAAACAG	1080
Qy	1081	GTCACTGACTCCCTGGTGCACCTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Db	1081	GTCACTGACTCCCTGGTGCACCTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTTAA	1200
Db	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTTAA	1200
Qy	1201	GGAGAAGAACGCTGTTATTTATGTTTAATCAATCAGAAATTCCTCACTGAGAAAGTTAAAGAA	1260
Db	1201	GGAGAAGAACGCTGTTATTTATGTTTAATCAATCAGAAATTCCTCACTGAGAAAGTTAAAGAA	1260
Qy	1261	ATTGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC	1320
Db	1261	ATTGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC	1320
Qy	1321	CTCAGCAATGGATGCCCTGGGTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATTG	1380
Db	1321	CTCAGCAATGGATGCCCTGGGTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATTG	1380
Qy	1381	TTACTCTCTTTGGACCCCTGTAICTTTAACTCCTCTGTTAAGTTTGTCTCTCCAGAATT	1440
Db	1381	TTACTCTCTTTGGACCCCTGTAICTTTAACTCCTCTGTTAAGTTTGTCTCTCCAGAATT	1440
Qy	1441	GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA	1481
Db	1441	GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA	1481
RESULT 2			
LOCUS	AR344387	1481 bp	linear
DEFINITION	Sequence 105 from patent US 6582703.		
ACCESSION	AR344387		
VERSION	AR344387.1	GI:33740328	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1481)		
AUTHORS	Perron, H., Beseme, F., Bedin, F., Paranhos-Baccalla, G., Komurian-Pradel, F., Jolivet-Reynaud, C. and Mandrand, B.		
TITLE	Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting		
JOURNAL	Patent: US 6582703-A 105 24-JUN-2003;		
FEATURES	Location/Qualifiers		
source	1..1481		
	/organism="unknown"		
	/mol_type="mRNA"		
ORIGIN			
Query Match	100.0%;	Score 1481;	DB 6; Length 1481;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1481;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0
Qy	1	ATGGGCGCTCCCTTATCATACATTTTCTCTTACTGTTCTTCTTACCCCTTTGCTCTCTCACT	60
Db	1	ATGGGCGCTCCCTTATCATACATTTTCTCTTACTGTTCTTCTTACCCCTTTGCTCTCTCACT	60
Qy	61	GCACCCCTCCATGCTGCTGTACACACAGTAGTCCCTTTACCAAGAGTTTCTATGAAGA	120
Db	61	GCACCCCTCCATGCTGCTGTACACACAGTAGTCCCTTTACCAAGAGTTTCTATGAAGA	120
Qy	121	ACGCGGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTCTTAAGGAAACTCT	180


```
Db 361 CAAGTCAGGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTATACATGAACCCCTCGGTACC 480
Db 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTATACATGAACCCCTCGGTACC 480
Qy 481 CATACTCGCTGGTGAGCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 481 CATACTCGCTGGTGAGCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCC 540
Qy 541 CAAAACCCCTACTAACTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
Db 541 CAAAACCCCTACTAACTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
Qy 601 CTGTTCCTGAAACAATGGAACAACCTTCAGCAGAGAAATAACACACCTCCGTTTAGTA 660
Db 601 CTGTTCCTGAAACAATGGAACAACCTTCAGCAGAGAAATAACACACCTCCGTTTAGTA 660
Qy 661 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAACCTCACCTGTCTAAATTT 720
Db 661 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAACCTCACCTGTCTAAATTT 720
Qy 721 AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACCTCCCAACGA 780
Db 721 AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACCTCCCAACGA 780
Qy 781 ATAGTCGTGCTACCTCAGAAATATTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 781 ATAGTCGTGCTACCTCAGAAATATTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 841 AATGGCTCTTCAGAACTATGCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC 900
Db 841 AATGGCTCTTCAGAACTATGCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC 900
Qy 901 ACTGAAACAAGATTTATACAACTCATGTCTGTAACCTTAAGCCCCCAACAACAAAGAGTACCATT 960
Db 901 ACTGAAACAAGATTTATACAACTCATGTCTGTAACCTTAAGCCCCCAACAACAAAGAGTACCATT 960
Qy 961 CTTCTCTTTGTTATCAGAGAGAGGTCTAGGAGACTAGGTACTGGCAATGGCGATATC 1020
Db 961 CTTCTCTTTGTTATCAGAGAGAGGTCTAGGAGACTAGGTACTGGCAATGGCGATATC 1020
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTCACATGGAACAG 1080
Db 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTCACATGGAACAG 1080
Qy 1081 GTCAGTACCTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT 1140
Db 1081 GTCAGTACCTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT 1140
Qy 1141 CAAAATCGAAGAGCTTTAGACTTGCTTAAACCGCCAAAGAGGGGGAACCTGTTATTTTA 1200
Db 1141 CAAAATCGAAGAGCTTTAGACTTGCTTAAACCGCCAAAGAGGGGGAACCTGTTATTTTA 1200
Qy 1201 GGAGAGAACGCTGTTATTTATGTTTAACTCAATCAGAAATGTCATCAGAAAGTTAAAGAA 1260
Db 1201 GGAGAGAACGCTGTTATTTATGTTTAACTCAATCAGAAATGTCATCAGAAAGTTAAAGAA 1260
Qy 1261 ATTCTGAGATCGAATAACAATAGAGCAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 1320
Db 1261 ATTCTGAGATCGAATAACAATAGAGCAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 1320
Qy 1321 CTAGCCAAATGGATGCGCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAAATTTG 1380
Db 1321 CTAGCCAAATGGATGCGCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAAATTTG 1380
Qy 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCCTTGTGTTAAGTTGTCTCTCCAGAAAT 1440
Db 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCCTTGTGTTAAGTTGTCTCTCCAGAAAT 1440
Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
```

RESULT 4

AF127228

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

misc_feature

misc_difference

ORIGIN

Query Match

Best Local Similarity

Matches 1481; Conservative

Score 1481; DB 14; Length 1932;

Pred. No. 0; Mismatches

Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTCTCTTACCCCTTTGCTCTCACT 60

Db 452 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTCTCTTACCCCTTTGCTCTCACT 511

QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 120

Db 512 GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 571

QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180

Db 572 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 631

QY 181 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Db 632 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 691

QY 241 CATGCAAACTCATATTATGGACAGGAAATGATTAATCTTAGTGTCTCTGAGGACTT 300

Db 692 CATGCAAACTCATATTATGGACAGGAAATGATTAATCTTAGTGTCTCTGAGGACTT 751

QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGTATCTGTGATGGGTGGAAAT 360

Db 752 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGTATCTGTGATGGGTGGAAAT 811

AF127228 1932 bp mRNA linear VRL 11-AUG-1999
Multiple sclerosis associated retrovirus element clone CL15
env-like mRNA sequence.

AF127228
AF127228.1 GI:5726293

Multiple sclerosis associated retrovirus element
Multiple sclerosis associated retrovirus element
Viruses; Retroviral viruses; Retroviridae.

1 (bases 1 to 1932)

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F.,
Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E.,
Mallet, F., Mandrand, B. and Perron, H.
Molecular cloning and characterization of MSRV-related sequences
associated with retrovirus-like particles

Virolgy 260 (1), 1-9 (1999)

9935590

10405350

2 (bases 1 to 1932)

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F.,
Ounanian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E.,
Mallet, F., Mandrand, B. and Perron, H.

Direct Submission

Submitted (10-FEB-1999) UMR103 CNRS, bioMerieux, 46, allée

D'Italie, Lyon 69007, France

Location/Qualifiers

1. .1932

/organism="Multiple sclerosis associated retrovirus

element"

/mol_type="mRNA"

/db_xref="taxon:89382"

/clones="CL15"

452. .1932

/note="similar to env but interrupted by premature stop

codon; includes signal peptide region, and partial SU and

TM regions"

misc_difference 568

/note="in clones CL15-2 and CL37-14; results in a

premature stop codon to Trp amino acid translation"

/replace="g"

Query Match 100.0%; Score 1481; DB 14; Length 1932;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches

Matches 1481; Conservative 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTCTCTTACCCCTTTGCTCTCACT 60

Db 452 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTCTCTTACCCCTTTGCTCTCACT 511

QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 120

Db 512 GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 571

QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180

Db 572 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 631

QY 181 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Db 632 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 691

QY 241 CATGCAAACTCATATTATGGACAGGAAATGATTAATCTTAGTGTCTCTGAGGACTT 300

Db 692 CATGCAAACTCATATTATGGACAGGAAATGATTAATCTTAGTGTCTCTGAGGACTT 751

QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGTATCTGTGATGGGTGGAAAT 360

Db 752 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGTATCTGTGATGGGTGGAAAT 811

repeat_region	1143. .1456 /note="AluSx repeat: matches 1. .312 of consensus"	misc_feature	/note="match: GSS: Em:AQ054794 Em:AQ823559"
repeat_region	1457. .2938 /note="L1MEC repeat: matches 356. .1897 of consensus"	repeat_region	16067. .16294 /note="match: GSS: Em:AQ457003"
repeat_region	2922. .3739 /note="L1PA2 repeat: matches 5329. .6146 of consensus"	repeat_region	16923. .16998 /note="38 copies 2 mer at 71% conserved"
repeat_region	3844. .4087 /note="L1M4 repeat: matches 1. .251 of consensus"	repeat_region	17107. .17166 /note="L2 repeat: matches 2688. .2750 of consensus"
repeat_region	4092. .4231 /note="L1M4 repeat: matches 612. .756 of consensus"	repeat_region	17175. .17755 /note="L2 repeat: matches 1244. .1898 of consensus"
repeat_region	4373. .4559 /note="L1M4C repeat: matches 971. .1152 of consensus"	repeat_region	17839. .18106 /note="MER70A repeat: matches 148. .418 of consensus"
repeat_region	4586. .4619 /note="17 copies 2 mer aa 100% conserved"	repeat_region	18213. .18318 /note="L2 repeat: matches 1128. .1247 of consensus"
repeat_region	4623. .4964 /note="MLT1A1 repeat: matches 3. .333 of consensus"	repeat_region	18706. .18811 /note="L2 repeat: matches 2631. .2750 of consensus"
repeat_region	4989. .5149 /note="L1M8 repeat: matches 6120. .6288 of consensus"	repeat_region	18812. .19057 /note="AluSx repeat: matches 1. .297 of consensus"
repeat_region	5152. .5602 /note="L1PA15 repeat: matches 5325. .5779 of consensus"	repeat_region	19058. .19492 /note="L2 repeat: matches 2213. .2647 of consensus"
repeat_region	5606. .5916 /note="AluSx repeat: matches 1. .313 of consensus"	misc_feature	complement(19352. .19843) /note="match: GSS: Em:AQ340313"
repeat_region	5917. .6014 /note="49 copies 2 mer ga 75% conserved"	repeat_region	19556. .19872 /note="L1R17 repeat: matches 1. .326 of consensus"
repeat_region	6018. .6409 /note="L1PA15 repeat: matches 5769. .6157 of consensus"	repeat_region	19873. .21594 /note="HERV17 repeat: matches 6812. .8523 of consensus"
repeat_region	6566. .6750 /note="L1M4 repeat: matches 2857. .3036 of consensus"	repeat_region	21587. .22280 /note="HERV17 repeat: matches 4301. .4980 of consensus"
repeat_region	6751. .7063 /note="AluJo repeat: matches 1. .307 of consensus"	repeat_region	22298. .22742 /note="L1MB5 repeat: matches 5623. .6135 of consensus"
repeat_region	7064. .8097 /note="L1M4 repeat: matches 3036. .4105 of consensus"	repeat_region	22743. .23048 /note="AluJ repeat: matches 1. .306 of consensus"
repeat_region	8099. .8419 /note="AluSx repeat: matches 1. .306 of consensus"	repeat_region	23049. .23286 /note="L1MB5 repeat: matches 5403. .5623 of consensus"
repeat_region	8420. .8487 /note="L1M4C repeat: matches 5582. .5648 of consensus"	repeat_region	23285. .23361 /note="L1M4 repeat: matches 5006. .5082 of consensus"
repeat_region	10039. .10396 /note="THB1B repeat: matches 1. .364 of consensus"	repeat_region	23364. .23429 /note="L1M4 repeat: matches 4019. .4084 of consensus"
repeat_region	10519. .10807 /note="AluSx repeat: matches 3. .297 of consensus"	repeat_region	23429. .24305 /note="L1 repeat: matches 3097. .4043 of consensus"
repeat_region	11343. .11416 /note="37 copies 2 mer tt 67% conserved"	repeat_region	24306. .24593 /note="AluJo repeat: matches 1. .293 of consensus"
repeat_region	11448. .11983 /note="L1MD3 repeat: matches 7162. .7739 of consensus"	repeat_region	24594. .25109 /note="L1 repeat: matches 2600. .3097 of consensus"
repeat_region	12030. .12337 /note="AluSx repeat: matches 1. .309 of consensus"	repeat_region	25110. .25436 /note="L1M3 repeat: matches 5101. .5440 of consensus"
repeat_region	12995. .13077 /note="MIR repeat: matches 103. .186 of consensus"	repeat_region	25437. .25742 /note="AluSx repeat: matches 1. .307 of consensus"
repeat_region	13110. .13222 /note="FLAM C repeat: matches 2. .124 of consensus"	repeat_region	25779. .25961 /note="MIR repeat: matches 55. .260 of consensus"
repeat_region	13228. .13518 /note="L1MEC repeat: matches 2131. .2414 of consensus"	repeat_region	25962. .26059 /note="MIR repeat: matches 1. .89 of consensus"
repeat_region	13519. .13853 /note="MER7A repeat: matches 1. .346 of consensus"	repeat_region	26135. .26328 /note="MLT1A1 repeat: matches 144. .352 of consensus"
repeat_region	13854. .14039 /note="L1MEC repeat: matches 1951. .2131 of consensus"	repeat_region	26554. .26697 /note="MLT1A1 repeat: matches 1. .149 of consensus"
repeat_region	14040. .14254 /note="MER20 repeat: matches 1. .218 of consensus"	repeat_region	26711. .26936 /note="L1PA9 repeat: matches 5939. .6162 of consensus"
repeat_region	14255. .14807 /note="L1MEC repeat: matches 1686. .1951 of consensus"	repeat_region	26981. .27154 /note="L1M4C repeat: matches 5620. .5805 of consensus"
repeat_region	14808. .15105 /note="AluJb repeat: matches 1. .311 of consensus"	repeat_region	27182. .27307 /note="L1P4 repeat: matches 5454. .5584 of consensus"
repeat_region	15106. .15492 /note="L1MEC repeat: matches 1309. .1686 of consensus"		
repeat_region	15496. .15822 /note="L1MB5 repeat: matches 5823. .6167 of consensus"		
repeat_region	15824. .15920 /note="L1MEC repeat: matches 1044. .1140 of consensus"		
misc_feature	15885. .16086 /note="match: STS: Em:G34444 Em:L14259"		
misc_feature	15920. .16082 /note="match: GSS: Em:AQ769824"		
misc_feature	complement(15927. .16073)		
		Query Match	99.1%; Score 1468.2; DB 9; Length 162579;
		Best Local Similarity	99.5%; Pred. No. 0;
		Matches 1473; Conservative	0; Mismatches 8; Indels 0; Gaps 0;
Qy	1 ATGGCCCTCCCTATCATATCTTTCTCTTTACTGTTCTTTACCCCTTTCCCTTACCAAGAGTTTCTATGAAGA 120		
Db	21559 ATGGCCCTCCCTATCATATCTTTCTCTTTACTGTTCTTTACCCCTTTCCCTTACCAAGAGTTTCTATGAAGA 21400		
Qy	61 GCACCCCTCCCTATCATATCTTTCTCTTTACTGTTCTTTACCCCTTTCCCTTACCAAGAGTTTCTATGAAGA 120		
Db	21499 GCACCCCTCCCTATCATATCTTTCTCTTTACTGTTCTTTACCCCTTTCCCTTACCAAGAGTTTCTATGAAGA 21400		

QY 121 ACGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 180
| | | | |
Db 21439 ACGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 21380
| | | | |
QY 181 ACCTTCACCTGCCACACCCATATATGCCCGCAACTGCTATACTCTGCGCACTCTTTGCATG 240
| | | | |
Db 21379 ACCTTCACCTGCCACACCCATATATGCCCGCAACTGCTATACTCTGCGCACTCTTTGCATG 21320
| | | | |
QY 241 CATGCAAACTACTCAATTATTGGACAGGGAATGATTAATCTCTAGTTGCTCTGGAGACTT 300
| | | | |
Db 21319 CATGCAAACTACTCAATTATTGGACAGGGAATGATTAATCTCTAGTTGCTCTGGAGACTT 21260
| | | | |
QY 301 GGAGCCACTGTCTGTTGGACTTACTTACCCATPACCAGTATGTCATGATGGGGTGAAT 360
| | | | |
Db 21259 GGAGCCACTGTCTGTTGGACTTACTTACCCATPACCAGTATGTCATGATGGGGTGAAT 21200
| | | | |
QY 361 CRAAGTTCAGGCAAGAGAAAACAAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
| | | | |
Db 21199 CRAAGTTCAGGCAAGAGAAAACAAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 21140
| | | | |
QY 421 AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
| | | | |
Db 21139 AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 21080
| | | | |
QY 481 CATACTCGCTGTGTGAGCTTATTAATACCACTCCTCCTCGGCTCCATGAGGTCTCAGCC 540
| | | | |
Db 21079 CATACTCGCTGTGTGAGCTTATTAATACCACTCCTCCTCGGCTCCATGAGGTCTCAGCC 21020
| | | | |
QY 541 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAAGGCATACATTTCAATC 600
| | | | |
Db 21019 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAAGGCATACATTTCAATC 20960
| | | | |
QY 601 CTGTGTTCTGAAACAATGGAAACAATCTCAGACAGAAATAAACAACCTTCCGTTTAGTA 660
| | | | |
Db 20959 CTGTGTTCTGAAACAATGGAAACAATCTCAGACAGAAATAAACAACCTTCCGTTTAGTA 20900
| | | | |
QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCTACCTCAAACTCTCCTCTGTGAAATTT 720
| | | | |
Db 20899 GGACCTCTGTTTCCAACTCTGGAATAAACCCTACCTCAAACTCTCCTCTGTGAAATTT 20840
| | | | |
QY 721 AGCAATACTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
| | | | |
Db 20839 AGCAATACTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 20780
| | | | |
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 840
| | | | |
Db 20779 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 20720
| | | | |
QY 841 AATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
| | | | |
Db 20719 AATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20660
| | | | |
QY 901 ACTGAAACAAGATTTATCAATCATGTCTGTAACCTTAAGCCCAACAAAGAGTACCCATT 960
| | | | |
Db 20659 ACTGAAACAAGATTTATCAATCATGTCTGTAACCTTAAGCCCAACAAAGAGTACCCATT 20600
| | | | |
QY 961 CTTCCCTTTTGTATCAGACAGGAGTCTAGGCAGACTAGGTACTGGCATTTGGCAGTATC 1020
| | | | |
Db 20599 CTTCCCTTTTGTATCAGACAGGAGTCTAGGCAGACTAGGTACTGGCATTTGGCAGTATC 20540
| | | | |
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTGACATGGACAG 1080
| | | | |
Db 20539 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTGACATGGACAG 20480
| | | | |
QY 1081 GTCACTGACTCCCTGGTCACTTTCGAAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTT 1140
| | | | |
Db 20479 GTCACTGACTCCCTGGTCACTTTCGAAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTT 20420
| | | | |
QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTCTGTTTATTTT 1200
| | | | |
Db 20419 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTCTGTTTATTTT 20360
| | | | |

QY 1201 GGAGAGAAACGCTGTATTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGCTTAAAGAA 1260
| | | | |
Db 20359 GGAGAGAAAGTGTGTATTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGCTTAAAGAA 20300
| | | | |
QY 1261 ATTCAGATCGAATCAATGATAGACAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 1320
| | | | |
Db 20299 ATTCAGATCGAATCAATGATAGACAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 20240
| | | | |
QY 1321 CTCAGCAATGATGATCCCTGGGTTCTCCCTCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
| | | | |
Db 20239 CTCAGCAATGATGATCCCTGGGTTCTCCCTCTTCTTAGGACCTCTAGCAGCTCTAATATTG 20180
| | | | |
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTTTAACTTTGTCTCTTCCAGAAAT 1440
| | | | |
Db 20179 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTTTAACTTTGTCTCTTCCAGAAAT 20120
| | | | |
QY 1441 GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAAACCCCA 1481
| | | | |
Db 20119 GAAGCTGTAAAGCTACAAATGGTTCTTCAATGGAGCCCA 20079
| | | | |
RESULT 6
QY19381
LOCUS Sequence 5315 from Patent WO02068579. linear PAT 03-FEB-2004
DEFINITION QY19381
ACCESSION QY19381
VERSION QY19381.1 GI:42280238
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter, C. J., Adams, M. C., Li, P. W. and Myers, E. W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 5315 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Source Location/Qualifiers
1..2055
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.0%; Score 1466.6; DB 6; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1472; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATACTTTTCTTTACTTGTCTCTTACCCCTTTTCGCTCTCACT 60
| | | | |
Db 36 ATGGCCCTCCCTTATCATATTTTCTTTACTTGTCTCTTACCCCTTTTCGCTCTCACT 95
| | | | |
QY 61 GCACCCCTCCATGCTGTGTACACAGTAGTCCCTTTACCAAGAGTTTCTATGAAGA 120
| | | | |
Db 96 GCACCCCTCCATGCTGTGTACACAGTAGTCCCTTTACCAAGAGTTTCTATGAAGA 155
| | | | |
QY 121 ACGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 180
| | | | |
Db 156 ACGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 215
| | | | |
QY 181 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGTCTATAAATCTGTCACCTCTTTGATG 240
| | | | |
Db 216 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGTCTATAAATCTGTCACCTCTTTGATG 275
| | | | |
QY 241 CATGCAAACTACTCAATTATTGGACAGGGAATGATTAATCTTAGTTGCTCTGGAGGACTT 300
| | | | |
Db 276 CATGCAAACTACTCAATTATTGGACAGGGAATGATTAATCTTAGTTGCTCTGGAGGACTT 335
| | | | |
QY 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCATACCAAGTATGTCATGATGGGGTGAAT 360
| | | | |
Db 336 GGAGCCACTGTCTGTTGGACTTACTTCAACCATACCAAGTATGTCATGATGGGGTGAAT 395
| | | | |

QY	1441	GAAGCTGTAAAGCTACAGATGGTCTTTACAAATGGAACCCCA	1481
Db	1476	GAAGCTGTAAAGCTACAAATGGTCTTTCAAAATGGAGCCCA	1516
RESULT 7			
BD252040			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
OS Homo sapiens (human)			
PN JP 2002539804-A/1			
PD 26-NOV-2002			
PF 20-MAR-2000 JP 2000607009			
PR 19-MAR-1999 FR 99/03622,28-OCT-1999 FR 99/13755 PI			
PR HERVE PERRON,MONIQUE LAFONT			
PC C12N15/09,A61K31/711,A61K38/00,A61K39/395,A61K45/00,A61K48/00,			
PC A61P25/00,			
PC C07K14/47,C07K16/18,C12N7/00,C12Q1/02,C12Q1/68,G01N27/447,PC			
G01N33/15,			
PC			
G01N33/50,G01N33/569,G01N33/577//C12P21/08,(C12N7/00,C12R1:93),PC			
(C12Q1/02,C12R1:93),(C12Q1/68,C12R1:93),C12N15/00,G01N27/26,PC			
A61K37/02			
CC Method for detecting superantigen activity in a biological sample			
FH Key			
FT source			
FT			
Location/Qualifiers			
1..1629			
/organism='Homo sapiens'			
/mol_type='genomic DNA'			
/db_xref='taxon:9606'			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1469; Conservative			
98.7%; Score 1461.8; DB 6; Length 1629;			
99.8%; Pred. No. 0;			
0; Mismatches 12; Indels 0; Gaps 0			
QY	1	ATGCCCTCCCTTATCATACTTTTCTTTTACTTGTTCTCTTACCCCTTCGCTCTCACT	60
Db	1	ATGCCCTCCCTTATCATACTTTTCTTTTACTTGTTCTCTTACCCCTTCGCTCTCACT	60
QY	61	GCACCCCTCCATGCTGTGTACCAACAGTAGTCCCTTACCAAGATTTCTATGAAGA	120
Db	61	GCACCCCTCCATGCTGTGTACCAACAGTAGTCCCTTACCAAGATTTCTATGAAGA	120
QY	121	ACGGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACTCC	180
Db	121	ACGGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACTCC	180
QY	181	ACCTTCACTGCCACACCCATATGCCCGCACTGTCTATAAATCTAGTTGCTTCGATG	240
Db	181	ACCTTCACTGCCACACCCATATGCCCGCACTGTCTATAAATCTAGTTGCTTCGATG	240
QY	241	CATGCAAAATCTCATTTATGGACAGGAAATCATTAATCTAGTTGCTTCGAGGACTT	300
Db	241	CATGCAAAATCTCATTTATGGACAGGAAATCATTAATCTAGTTGCTTCGAGGACTT	300
QY	301	GGAGCCACTGTCTGTGGACTTACTTCAACCCATACCAAGTATGTCTGATGGGGTGAATT	360

Db	601	CTGTTCCTGAAACAATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGCTTTTAGTGA	660
Qy	661	GGACCTCTGTGTTTCCAAATCTGGAAATAACCCATACCTCAAACTCACCCTGTCTGAATAATT	720
Db	661	GGACCTCTGTGTTTCCAAATCTGGAAATAACCCATACCTCAAACTCACCCTGTCTGAATAATT	720
Qy	721	AGCAATACTATAGACACAACACAGCTCCCAATGATCAGGTGGGTAAACACTCCCAACAGA	780
Db	721	AGCAATACTATAGACACAACACAGCTCCCAATGATCAGGTGGGTAAACACTCCCAACAGA	780
Qy	781	ATAGTCTGCCTACCCCTCAGGAATATTTTTGTCTGTGTACCTCAGCCCTATCATTTGTTTG	840
Db	781	ATAGTCTGCCTACCCCTCAGGAATATTTTTGTCTGTGTACCTCAGCCCTATCATTTGTTTG	840
Qy	841	AATGGCTCTTCAGAAATATTTTTGTCTGTGTACCTCAGCCCTATCATTTGTTTG	900
Db	841	AATGGCTCTTCAGAAATATTTTTGTCTGTGTACCTCAGCCCTATCATTTGTTTG	900
Qy	901	ACTGAAACAAGATTTATACAATCATGTCTGTAAGCCCAACAAAAAGAGTACCATT	960
Db	901	ACTGAAACAAGATTTATACAATCATGTCTGTAAGCCCAACAAAAAGAGTACCATT	960
Qy	961	CTTCCTTTTGTATTCAGAGCAGAGTGCTTAGGCAGACTAGGTACTGGCATTGGCAGTATC	1020
Db	961	CTTCCTTTTGTATTCAGAGCAGAGTGCTTAGGCAGACTAGGTACTGGCATTGGCAGTATC	1020
Qy	1021	ACAACTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAATGGTGACATGGAACAG	1080
Db	1021	ACAACTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAATGGTGACATGGAACAG	1080
Qy	1081	GTCACTGACTCCCTGGTCACTTAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Db	1081	GTCACTGACTCCCTGGTCACTTAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Qy	1141	CAAAATCGAAGAGCTTTAGACTTCTAAACCGCAAAAGAGGGGAACCTGTTATTTTAA	1200
Db	1141	CAAAATCGAAGAGCTTTAGACTTCTAAACCGCAAAAGAGGGGAACCTGTTATTTTAA	1200
Qy	1201	GGAGAAGAACCTGTTATTATGTTTAACTCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA	1260
Db	1201	GGAGAAGAACCTGTTATTATGTTTAACTCAATCCAGAATTGTCACTGAGAAAGTTAAAGAA	1260
Qy	1261	ATTCGAGATCGAATACAATGTAGACGACAGAGAGCTTCAAAACACCGAACCTGGGGCTC	1320
Db	1261	ATTCGAGATCGAATACAATGTAGACGACAGAGAGCTTCAAAACACCGAACCTGGGGCTC	1320
Qy	1321	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAAATATTG	1380
Db	1321	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAAATATTG	1380
Qy	1381	TTACTCTCTTTTGGACCTGTATCTTTAACTCCCTTGTAAAGTTTGTCTTCCAGAAATT	1440
Db	1381	TTACTCTCTTTTGGACCTGTATCTTTAACTCCCTTGTAAAGTTTGTCTTCCAGAAATT	1440
Qy	1441	GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA	1481
Db	1441	GAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCCA	1481

RESULT 9	AF331500	AF331500	1629 bp	linear	VRL 21-NOV-2001
LOCUS	AF331500	AF331500	1629 bp	linear	VRL 21-NOV-2001
DEFINITION	AF331500	AF331500	Multiple sclerosis associated retrovirus element clone pV14	Multiple sclerosis associated retrovirus element clone pV14	Multiple sclerosis associated retrovirus element clone pV14
ACCESSION	AF331500	AF331500	Multiple sclerosis associated retrovirus element	Multiple sclerosis associated retrovirus element	Multiple sclerosis associated retrovirus element
VERSION	AF331500.1	AF331500.1	GI:13310190	GI:13310190	GI:13310190
KEYWORDS	AF331500.1	AF331500.1	GI:13310190	GI:13310190	GI:13310190
SOURCE	AF331500.1	AF331500.1	GI:13310190	GI:13310190	GI:13310190
ORGANISM	AF331500.1	AF331500.1	GI:13310190	GI:13310190	GI:13310190
REFERENCE	AF331500.1	AF331500.1	GI:13310190	GI:13310190	GI:13310190
AUTHORS	AF331500.1	AF331500.1	GI:13310190	GI:13310190	GI:13310190

TITLE	Camelo, S., Dumon, A., Jolivet-Reynaud, C., Marcel, F., Souillet, Y., Borel, S., Gebuhrer, L., Santoro, L., Marcel, S., Seigneurin, J. M., Marche, P. N. and Lafon, M. Multiple sclerosis retrovirus particles and recombinant envelope trigger an abnormal immune response in vitro, by inducing polyclonal Vbeta16 T-lymphocyte activation Virology 287 (2), 321-332 (2001)
JOURNAL	21425392
MEDLINE	11531410
PUBMED	2 (bases 1 to 1629)
REFERENCE	Perron, H., Ouanian-Paraz, A., Marcel, F. and Souillet, Y. Direct Submission TITLE Submitted (21-DEC-2000) R&D, bioMerieux, Chemin de L'Orme, Marcy L'Etoile Location/Qualifiers 1. .1629 /organism="Multiple sclerosis associated retrovirus element" /mol_type="mRNA" /db_xref="taxon:89382" /clone="pV14" 1. .1629 /genes="env" 1. .1629 /genes="env" /function="displaying superantigen-like properties causing polyclonal activation of human Vbeta 16 T-cells (CD3+), independent from the T-cell receptor antigen specificity" /notes="surface and transmembrane regions" /codon_start=1 /product="recombinant envelope protein" /protein_id="AAK18189.1" /db_xref="GI:13310191" /translations="MALPYHTFLFTVLLPPFALTAPPCCCTSSSPYOEFFLWRPLP GNIDAPSYRLSKSGNSTFTAHMPRNCYNYSATLCMHANTHYWTGKMINPSCPGGLGA TCWTFYTHFTSMDSNGTIOGAAREKOVKEAISOLTRGHSTPSPYKGLVLSKLEHTET HTRLVSLNLTITRLRHEVSNQNTFCWCLPLHFRPVISIPVPEQNNPSTENITTSV LVGLPVSNLBETHTSNLTCPKPSNTIDTSSQCIKRWTPPTKIVCLPSGIFVFCGTSIA YHCLNGSSSEMCFSLFVLPQNTDVTQEDLVNHYVPPHKNRVILFPTVRAGVLLGK TG3GSIITTSQFYKLSQINGEMQVDSLTVLQDLNLSAAVLVQNRRLDOLLK RGGTCLPLGBERQYVQNSRIVTEKVKETRDRIQCRAEELQNTERRGLGSLSQMWPWTLIP FLGAPLAEITFLLPGCINFLKPVSSRIEAVGLQVLQMEPQMSMTKIYRGPDLR PARICSDNVIDEYTPPEEISTAOPLLHNSVGS"
FEATURES	
source	
gene	
CDs	

ORIGIN	Query Match	98.7%	Score 1461.8	DB 14	Length 1629
	Best Local Similarity	99.2%	Pred. No. 0		
	Matches 1469	Conservative 0	Mismatches 12	Indels 0	Gaps 0
QY	1	ATGGCCCTCCCTTATCATCTTTTCTCTTTACTGTCTTACCCCTTTGGCTCTCACT	60		
Db	1	ATGGCCCTCCCTTATCATCTTTTCTCTTTACTGTCTTACCCCTTTGGCTCTCACT	60		
QY	61	GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA	120		
Db	61	GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA	120		
QY	121	ACGGGGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAAATCC	180		
Db	121	ACGGGGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAAATCC	180		
QY	181	ACCTTCAGTCGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGGCATG	240		
Db	181	ACCTTCAGTCGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGGCATG	240		
QY	241	CATGCAAAATCACTCAITATTGGACAGGGAAAAATGATTAATCTAGTTGTCCTGGAGGACTT	300		
Db	241	CATGCAAAATCACTCAITATTGGACAGGGAAAAATGATTAATCTAGTTGTCCTGGAGGACTT	300		
QY	301	GGAGCCATGTCTGTGGACTTACTTACCCCATACCAAGTAGTCTGATGGGGGTGGAATT	360		
Db	301	GGAGCCATGTCTGTGGACTTACTTACCCCATACCAAGTAGTCTGATGGGGGTGGAATT	360		

361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCCGGGACAT 420
 Db
 361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCCGGGACAT 420
 Qy
 421 AGCACCCCTAGCCCTTACAAAGGACTAGTCTCTCAAAATACATGAAACCCCTCCGTACC 480
 Qy
 421 AGCACCCCTAGCCCTTACAAAGGACTAGTCTCTCAAAATACATGAAACCCCTCCGTACC 480
 Db
 481 CATACTCGCTGGTGAGCCTATTATATACACACCTCACTCGGTCCATGAGGTCTCAGCC 540
 Qy
 481 CATACTCGCTGGTGAGCCTATTATATACACACCTCACTCGGTCCATGAGGTCTCAGCC 540
 Db
 541 CAATAACCTTACTTACTGTTGGATGTGCTCCCTCGCACTTTCAGGCGCATATTTCAATC 600
 Qy
 541 CAATAACCTTACTTACTGTTGGATGTGCTCCCTCGCACTTTCAGGCGCATATTTCAATC 600
 Db
 601 CTTGTTCTCTGAAACAATGGAACAATCTCAGACAGAAATAAACACCACTTCCGTTTGTAGTA 660
 Qy
 601 CTTGTTCTCTGAAACAATGGAACAATCTCAGACAGAAATAAACACCACTTCCGTTTGTAGTA 660
 Db
 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAACACCTCACTGTGTAAATTT 720
 Qy
 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAACACCTCACTGTGTAAATTT 720
 Db
 721 AGCAATACTATAGACACACAGCTCCCAATGATCAGTGGGTAAACACCTCCACACAGA 780
 Qy
 721 AGCAATACTATAGACACACAGCTCCCAATGATCAGTGGGTAAACACCTCCACACAGA 780
 Db
 781 ATAGTCTGCTTACCTCAGGAATATTTTTGTGTGTGTGCTACCTCAGCCTATCATTTGTTG 840
 Qy
 781 ATAGTCTGCTTACCTCAGGAATATTTTTGTGTGTGTGCTACCTCAGCCTATCATTTGTTG 840
 Db
 841 AATGGCTCTCAGGAATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 Qy
 841 AATGGCTCTCAGGAATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 Db
 901 ACTGAACAAGATTTATCAATCATCTGCTACCTAAGCCCAACAAAGAGTACCCATT 960
 Qy
 901 ACTGAACAAGATTTATCAATCATCTGCTACCTAAGCCCAACAAAGAGTACCCATT 960
 Db
 961 CTTCTCTTTGTTATCAGACGAGGAGTCTAGGACAGACTAGTACTGCGCAATGGCAGTATC 1020
 Qy
 961 CTTCTCTTTGTTATCAGACGAGGAGTCTAGGACAGACTAGTACTGCGCAATGGCAGTATC 1020
 Db
 1021 ACACCTCTACTAGTCTTACTACAACTATCTCAAGAAATAATGTCATGACATGACACAG 1080
 Qy
 1021 ACACCTCTACTAGTCTTACTACAACTATCTCAAGAAATAATGTCATGACATGACACAG 1080
 Db
 1081 GTCACTGACTCCCTGGTCCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT 1140
 Qy
 1081 GTCACTGACTCCCTGGTCCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT 1140
 Db
 1141 CAAATCGAAGAGCTTTAGACTTCTAACCGCCAAAAGAGGGGAACTGTGTTATTTTAA 1200
 Qy
 1141 CAAATCGAAGAGCTTTAGACTTCTAACCGCCAAAAGAGGGGAACTGTGTTATTTTAA 1200
 Db
 1201 GGAGAAGAACCTGTTATTTATTAATCAATCCAGATTTGCTACTGAGAAAGTTAAAGAA 1260
 Qy
 1201 GGAGAAGAACCTGTTATTTATTAATCAATCCAGATTTGCTACTGAGAAAGTTAAAGAA 1260
 Db
 1261 ATTCGAGATCGAATACAAATGAGACGAGGAGCTTCAAAACACCGCAACCGTGGGGCTC 1320
 Qy
 1261 ATTCGAGATCGAATACAAATGAGACGAGGAGCTTCAAAACACCGCAACCGTGGGGCTC 1320
 Db
 1321 CTCAGCCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
 Qy
 1321 CTCAGCCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
 Db
 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTTGTAGTTGTCTCTCCAGATT 1440
 Qy
 1381 TTACTCTCTTTGGACCTGTATCTTTCAACTCTCTTGTAAAGTTGTCTCTCCAGATT 1440
 Db
 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 Qy

Db 1441 GAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCCA 1481
 RESULT 10
 AC023366
 LOCUS
 DEFINITION
 AC023366 172918 bp DNA linear HTG 02-MAR-2000
 Homo sapiens clone RP11-15N10, WORKING DRAFT SEQUENCE, 33 unordered
 pieces.
 AC023366 GI:7143443
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 172918)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 2 (bases 1 to 172918)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bada, F., Boguslavskiy, L.,
 Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
 Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cooke, P.,
 DeArallano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,
 Feneator, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., Landers, T., Largocque, K., Lehoczy, J., Levine, R.,
 Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
 McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J.,
 Meneus, L., Mihova, T., Miranda, C., Menga, V., Morrow, J., Naylor, J.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivari, T. M.,
 Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C.,
 Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A.,
 Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
 Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and
 Zody, M.
 Direct Submission
 Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 2, 2000 this sequence version replaced gi:6970363.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L3492
 Center clone name: 15 N.10
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 142569 bases at least Q40
 Consensus quality: 156863 bases at least Q30
 Consensus quality: 164107 bases at least Q20
 Insert size: 178000; agarose-fp
 Insert size: 169718; sum-of-contigs
 Quality coverage: 2.9 in Q20 bases; agarose-fp
 Quality coverage: 3.0 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 33 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

- * runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence
- * as soon as it is available and the accession number will
- * be preserved.

*	1	1266:	contig of 1266 bp in length
*	1267	1366:	gap of 100 bp
*	1367	2686:	contig of 1320 bp in length
*	2687	2786:	gap of 100 bp
*	2787	4288:	contig of 1502 bp in length
*	4289	4388:	gap of 100 bp
*	4389	5572:	contig of 1184 bp in length
*	5573	5672:	gap of 100 bp
*	5673	6812:	contig of 1140 bp in length
*	6813	6912:	gap of 100 bp
*	6913	7404:	contig of 492 bp in length
*	7405	7594:	gap of 100 bp
*	7595	8915:	contig of 1411 bp in length
*	8916	9015:	gap of 100 bp
*	9016	10200:	contig of 1185 bp in length
*	10201	10300:	gap of 100 bp
*	10301	12003:	contig of 1703 bp in length
*	12004	12103:	gap of 100 bp
*	12104	14038:	contig of 1935 bp in length
*	14039	14138:	gap of 100 bp
*	14139	16226:	contig of 2098 bp in length
*	16227	16326:	gap of 100 bp
*	16327	19115:	contig of 2789 bp in length
*	19116	19215:	gap of 100 bp
*	19216	21432:	contig of 2277 bp in length
*	21433	21532:	gap of 100 bp
*	21593	24794:	contig of 3202 bp in length
*	24795	24894:	gap of 100 bp
*	24895	27624:	contig of 2730 bp in length
*	27625	27724:	gap of 100 bp
*	27725	30780:	contig of 3056 bp in length
*	30781	30880:	gap of 100 bp
*	30881	34253:	contig of 3373 bp in length
*	34254	34353:	gap of 100 bp
*	34354	38260:	contig of 3907 bp in length
*	38261	38360:	gap of 100 bp
*	38361	42757:	contig of 4397 bp in length
*	42758	42857:	gap of 100 bp
*	42858	46232:	contig of 3375 bp in length
*	46233	46332:	gap of 100 bp
*	46333	51049:	contig of 4717 bp in length
*	51050	51149:	gap of 100 bp
*	51150	53373:	contig of 2224 bp in length
*	53374	53473:	gap of 100 bp
*	53474	58170:	contig of 4697 bp in length
*	58171	58270:	gap of 100 bp
*	58271	63398:	contig of 5128 bp in length
*	63399	63498:	gap of 100 bp
*	63499	63283:	contig of 5785 bp in length
*	63284	69383:	gap of 100 bp
*	69384	75177:	contig of 5794 bp in length
*	75178	75277:	gap of 100 bp
*	75278	80817:	contig of 5540 bp in length
*	80818	80917:	gap of 100 bp
*	80918	89553:	contig of 8636 bp in length
*	89554	89653:	gap of 100 bp
*	89654	101099:	contig of 11446 bp in length
*	101100	101199:	gap of 100 bp
*	101200	113135:	contig of 11936 bp in length
*	113136	113235:	gap of 100 bp
*	113236	134520:	contig of 19285 bp in length
*	134521	136260:	gap of 100 bp
*	136261	150646:	contig of 18026 bp in length
*	150647	150746:	gap of 100 bp
*	150747	172918:	contig of 22172 bp in length

FEATURES

```

/db_xref="taxon:9606"
/clone="RP11-15N10"
/clone_lib="RPC1-11 Human Male BAC"
1. .1266
/misc_feature
/notes="assembly_fragment"
1367. .2686
/misc_feature
/notes="assembly_fragment"
2787. .4288
/misc_feature
/notes="assembly_fragment"
4389. .5572
/misc_feature
/notes="assembly_fragment"
5673. .6812
/misc_feature
/notes="assembly_fragment"
6913. .7404
/misc_feature
/notes="assembly_fragment"
clone_end:T7
vector_side:right
7505. .8915
/misc_feature
/notes="assembly_fragment"
9016. .10200
/misc_feature
/notes="assembly_fragment"
10301. .12003
/misc_feature
/notes="assembly_fragment"
12104. .14038
/misc_feature
/notes="assembly_fragment"
14139. .16226
/misc_feature
/notes="assembly_fragment"
16327. .19115
/misc_feature
/notes="assembly_fragment"
19216. .21492
/misc_feature
/notes="assembly_fragment"
21593. .24794
/misc_feature
/notes="assembly_fragment"
24895. .27624
/misc_feature
/notes="assembly_fragment"
27725. .30780
/misc_feature
/notes="assembly_fragment"
30881. .34253
/misc_feature
/notes="assembly_fragment"
34354. .38260
/misc_feature
/notes="assembly_fragment"
38361. .42757
/misc_feature
/notes="assembly_fragment"
42858. .46232
/misc_feature
/notes="assembly_fragment"
46333. .51049
/misc_feature
/notes="assembly_fragment"
51150. .53373
/misc_feature
/notes="assembly_fragment"
clone_end:SP6
vector_side:right
53474. .58170
/misc_feature
/notes="assembly_fragment"
58271. .63398
/misc_feature
/notes="assembly_fragment"
63499. .69283
/misc_feature
/notes="assembly_fragment"
69384. .75177
/misc_feature
/notes="assembly_fragment"
75278. .80817
/misc_feature
/notes="assembly_fragment"

```

Query Match	98.7%	Score 1461.8	DB 2	Length 172918
Best Local Similarity	99.2%	Pred. No. 0		
Matches 1469	Conservative 0	Mismatches 12	Indels 0	Gaps 0
Qy 1	ATFGGCGCTCCCTTATCATACATCTTTCTCTTTACTGTCTCTTACCCCGCTTTTCGCTCTCACT	60		
Db 103234	ATFGGCGCTCCCTTATCATATTTTCTCTTTACTGTCTCTCAACCCCGCTTTTCGCTCTCACT	103293		
Qy 61	GCACCCCGCTCAGTGTGTGTGTACACACAGTAGTCTCCCTTTACCAAGAGTTTCTTATCAAGA	120		
Db 103294	GCACCGCTCAGTGTGTGTGTGTACACACAGTAGTCTCCCTTTACCAAGAGTTTCTTATAAGA	103353		

QY 121 ACGGGCTTCTGTGAATATTTGATGCCCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
 DB 103354 ACGGGCTTCTGTGAATATTTGATGCCCCCATCATATAGGAGTTTATCTAAGGAAACTCC 103413
 QY 181 ACTTCACTGCCACACCCATATGCCCGGAACTGCTATTAATCTGCGACTCTTTGCAATG 240
 DB 103414 ACTTCACTGCCACACCCATATGCCCGGAACTGCTATTAATCTGCGACTCTTTGCAATG 103473
 QY 241 CATGCAAACTACTCAATATTTGACAGGAAATGATTAATCTAGTTGCTCTGAGGAGACTT 300
 DB 103474 CATGCAAACTACTCAATATTTGACAGGAAATGATTAATCTAGTTGCTCTGAGGAGACTT 103533
 QY 301 GGAGCCACTGTCTGTGAGCTTACTTCAACCCATACAGATATGCTGATGCGGGGFGAAATT 360
 DB 103534 GGAGCCACTGTCTGTGAGCTTACTTCAACCCATACAGATATGCTGATGCGGGGFGAAATT 103593
 QY 361 CAAGGTGAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGACAT 420
 DB 103594 CAAGGTGAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGACAT 103653
 QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 DB 103654 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 103713
 QY 481 CATACTCGCCCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
 DB 103714 CATACTCGCCCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 103773
 QY 541 CAAACCCCTACTAATCTGTGAGTGTGCTCCCTCGCACTTTCAGGCGCATATCATTTCAATC 600
 DB 103774 CAAACCCCTACTAATCTGTGAGTGTGCTCCCTCGCACTTTCAGGCGCATATCATTTCAATC 103833
 QY 601 CCTGTTCTGAACTGGAACAACTTCAGACACAAATTAACACCACTTCGCTTTTAACTA 660
 DB 103834 CCTGTTCTGAACTGGAACAACTTCAGACACAAATTAACACCACTTCGCTTTTAACTA 103893
 QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAACCTCAGCTGTGTAATTT 720
 DB 103894 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAACCTCAGCTGTGTAATTT 103953
 QY 721 AGCAATACTATAGACACAACTGCAATGCAATGAGTGGGTAAACCTTCCACACGA 780
 DB 103954 AGCAATACTATAGACACAACTGCAATGCAATGAGTGGGTAAACCTTCCACACGA 104013
 QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGCTGAGTACCTCAGCTATCATTTGTTG 840
 DB 104014 ATAGTCTGCTACCTCAGGAATATTTTGTCTGCTGAGTACCTCAGCTATCATTTGTTG 104073
 QY 841 AATGGCTCTCAGAACTATATGCTGCTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
 DB 104074 AATGGCTCTCAGAACTATATGCTGCTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 104133
 QY 901 ACTGAACAAGATTTATCAATCATGTGCTACCTAAGCCCAACAAACAAAGAGTACCCATT 960
 DB 104134 ACTGAACAAGATTTATCAATCATGTGCTACCTAAGCCCAACAAACAAAGAGTACCCATT 104193
 QY 961 CTTCTTTTGTATCAGACGAGGAGTGTAGGAGACTAGTACTGCGATTGGCAGTATC 1020
 DB 104194 CTTCTTTTGTATCAGACGAGGAGTGTAGGAGACTAGTACTGCGATTGGCAGTATC 104253
 QY 1021 ACAACCTCTACTAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
 DB 104254 ACAACCTCTACTAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 104313
 QY 1081 GTCACTGATCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGAGTATGCTCTT 1140
 DB 104314 GTCACTGATCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGAGTATGCTCTT 104373
 QY 1141 CAAATCGAAGAGCTTTAGACTTGTACCCGCAAGAGGGGAACTGTTTATTTTAA 1200
 DB 104374 CAAATCGAAGAGCTTTAGACTTGTACCCGCAAGAGGGGAACTGTTTATTTTAA 104433
 QY 1201 GGAGAAGAACGCTGTTATTTATGTTAAATCAATCCAGAAATTGTCACTGAGAAAGTTTAAAGAA 1260

DB 104434 GGAGAAGAACGCTGTTATTTATGTTAAATCAATCCAGAAATTGTCTACTGAGAAAGTTAAAGAA 104493
 QY 1261 ATTCGAGATCGAATACAATGCTAGCAGAGAGAGCTTCAAAACACCGAACGCTGGGGCCTC 1320
 DB 104494 ATTCGAGATCGAATACAATGCTAGCAGAGAGAGCTTCAAAACACCGAACGCTGGGGCCTC 104553
 QY 1321 CTCAGCAATGGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAAATATTG 1380
 DB 104554 CTCAGCAATGGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAAATATTG 104613
 QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGAAATT 1440
 DB 104614 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGAAATT 104673
 QY 1441 GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481
 DB 104674 GAAGCTGTAAAGCTACAAATGGTCTTCAAAATGGAGCCCA 104714
 RESULT 11
 LOCUS AX478550 2074 bp DNA linear PAT 12-AUG-2002
 DEFINITION Sequence 4 from Patent: WO0248362.
 ACCESSION AX478550
 VERSION AX478550.1 GI:22217319
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ramkumar, J. and Arvizu, C.
 Embryogenesis associated proteins
 Patent: WO 0248362-A 4 20-JUN-2002;
 Incyte Genomics, Inc. (US)
 Location/Qualifiers
 source
 1. 2074
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7477736CB1"
 ORIGIN
 Query Match 93.8%; Score 1389.8; DB 6; Length 2074;
 Best Local Similarity 96.2%; Pred. No. 0;
 Matches 1424; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATCATACTTTTCTTTACTGTTCTTACCCCTTTTGGCTCTCACT 60
 DB 64 ATGGCCCTCCCTTATGTTATTTTCTTTACTGTTCTTCAACCCCTTTTCACTCACT 123
 QY 61 GCACCCCTCCCTGCTGTACACACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 120
 DB 124 GCACCCCTCCCTGCTGTACACACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 183
 QY 121 ACGCGGCTTCTGGAATATTTGATGCCCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
 DB 184 ATGGCGGTCTCCAGACATATTTGATGCCCCCATCGTATAGGAGTTTATCTAAGGAAACTCC 243
 QY 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATTAATCTGCCACTTTTTCATG 240
 DB 244 GCCTTACCGCCACACCCATATGCCCGCAACTGCTATTAATCTGCCACTTTTTCATG 303
 QY 241 CATGCAAACTACTCAATTTGACAGGAAATGATTAATCTAGTTGCTCTGAGGAGACTT 300
 DB 304 CATGCAAACTACTCAATTTGACAGGAAATGATTAATCTAGTTGCTCTGAGGAGACTT 363
 QY 301 GGAGCCACTGTCTGTTGAGCTTACTTCAACCCATACAGTATGCTGATGCGGGTGGAAATT 360
 DB 364 GGAGCCACTGTCTGTTGAGCTTACTTCAACCCATACCGGTATGCTGATGCGGGTGGAGTT 423
 QY 361 CAAGGTGAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGACAT 420

Db 1504 GAAGCTGTAAAGCTACAAATGGTTCCTCAATGAGGCCA 1544

RESULT 12
AY101589 10230 bp DNA linear PRI 11-FEB-2004
LOCUS Gorilla gorilla isolate 1 endogenous retrovirus ERV-W, ERVWEL
DEFINITION locus, allele B, complete sequence.
ACCESSION AY101589
VERSION AY101589.1 GI:37544413
KEYWORDS
SOURCE Gorilla gorilla (gorilla)
ORGANISM Gorilla gorilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Gorilla.

REFERENCE 1 (bases 1 to 10230)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVWEL is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826

REFERENCE 2 (bases 1 to 10230)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
Direct Submission
Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers
1..10230
/organism="Gorilla gorilla"
/mol_type="genomic DNA"
/isolate="J1"
/isolation_source="PBMC of gorilla 1"
/db_xref="taxon:9593"
/cell_type="PBMC"
/note="isolated by PCR; ERVWEL locus, allele B
endogenous_virus=ERV-W"
1..795
join(248..826,2894..9821)
/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..826,7595..9821)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..826,9258..9821)
/note="putative mRNA transcript 3; no coding region"
7822..9438
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA068168.1"
/db_xref="GI:37544414"

LTR
mRNA
mRNA
mRNA
CDS

ORIGIN
Query Match 90.6%; Score 1341.8; DB 9; Length 10230;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1 ATGCCCTCCCTTATCATCTTTCTTACTGTCTTCTTACCCCCCTTGCCTCTACT 60
DB 7822 ATGCCCTCCCTTATCATCTTTCTTACTGTCTTCTTACCCCCCTTGCCTCTACT 7881

Qy	61	GCACCCCTCCATGCTGCTGTACAAACAGTAGCTCCCTTTACAAAGAGTTTCTATGAAGA	120
Db	7882	GCACCCCTCCATGCGCGCTGTATGACAGTAGCTCCCTTTACCAAGAGTTTCTATGGAGA	7941
Qy	121	ACGGCGTTCCTGGAATATTGATGCGCCCATATAGGAGTTTATCTAAGGGAAACTCC	180
Db	7942	ATGCGCGTCTCGGAATATTGATGCGCCCATCGCATAGGAGTTTCTAAGGGAAACCCC	8001
Qy	181	ACCTTCACCTGCCACACCCATATGCCCAGCACTGCTATAACTCTGCCACTCTTTGCATG	240
Db	8002	ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG	8061
Qy	241	CATGCAAACTACTCATTTATTGGA CAGGAAATGATTAATCCTTAGTTGTCTCTGAGACATT	300
Db	8062	CATGCAAACTACTCATTTATTGGA CAGGAAATGATTAATCCTTAGTTGTCTCTGAGACATT	8121
Qy	301	GGAGCCACTGCTGTTGGACTTACTTACCACATACAGATATGCTCATGATGGGGGTGAAATT	360
Db	8122	GGAGTCACTGCTGTTGGACTTACTTACCACATACAGATATGCTCATGATGGGGGTGAGATT	8181
Qy	361	CAAGGTCAGCAAGAGAAAAAAGTAAAGGAAGCAATCTCCAACTGACCCCGGGACAT	420
Db	8182	CAAGATCAGCAGAGAAAAACATGTAAAGAGTAATCTCCCACTGACCCCGGTACAT	8241
Qy	421	AGCACCCCTAGCCCCCTACAAAGACTAGATTCTCTCAAACTACATGAAACCTCCCGTACC	480
Db	8242	AGCACCTTAGCCCCCTACAAAGACTAGATTCTCTCAAACTACATGAAACCTCCCGTACC	8301
Qy	481	CATACTGCGCTGTGAGCCTATTTTATACACCTCACTCGGCTCCATGAGGTCTCAGCC	540
Db	8302	CATACTGCGCTGTGAGCCTATTTTATACACCTCACTCGGCTCCATGAGGTCTCAGCC	8361
Qy	541	CAAAACCTACTTAACCTGCTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC	600
Db	8362	CAAAACCTACTTAACCTGCTGGATGTGCTCCCTCGCACTTCAGGCCATATGTTCAATC	8421
Qy	601	CCTGTTCTCTGAACAATGGAA CAACCTCAGCACAGAAATAAACACCACTTCGTTTGTAGTA	660
Db	8422	CCTGTACTCTGAAGAATGGAA CAACCTCAGCACAGAAATAAACACCACTTCGTTTGTAGTA	8481
Qy	661	GGACCTCTGTTTCCCAATCTGGAATAACCATACCTCAAACTCACTCTGTGTAAATTT	720
Db	8482	GGACCTCTGTTTCCCAATCTGGAATAACCATACCTCAAACTCACTCTGTGTAAATTT	8541
Qy	721	AGCAATACTATAGACACAAACAGCTCCCAATGCAATCAGTGGGTAAACACCTCCACACGA	780
Db	8542	AGCAATACTATAGACACAAACCACTCCCAATGCAATCAGTGGGTAACTCTCCACACAA	8601
Qy	781	ATAGTCTGCCCTACCTCAGGAATTTTTTGTCTGTGTACCTCAGGCTATCATTTGTTTG	840
Db	8602	ATAGTCTGCCCTACCTCAGGAATTTTTTGTCTGTGTACCTCAGGCTATCATTTGTTTG	8661
Qy	841	AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCCCCCTATGACCATCTAC	900
Db	8662	AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCCCCCTATGACCATCTAC	8721
Qy	901	ACTGAAACAAGATTTATACAATCATGTCTGTACTAAGCCCAACAAACAAAGAGTACCCATT	960
Db	8722	ACTGAAACAAGATTTATACAATATATGTGTATCTAAGCCCCCAACAAAGAGTACCCATT	8781
Qy	961	CTTCTCTTTTGTATCAGACGAGGAGTGCTAGGCAGATAGGTACTGCAATTTGGCAGTATC	1020
Db	8782	CTTCTCTTTTGTATATAGGACGAGGAGTGCTAGGTGCACTAGGTACTGCAATTTGGCGGTATC	8841
Qy	1021	ACAACCTCTACTCAGTTCTACTACAACATATCTCAAGNAATAAATGGTGTGACATGGAACAG	1080
Db	8842	ACAACCTCTACTCAGTTCTACTACAACATATCTCAAGNAATAAATGGGACATGGAACGG	8901
Qy	1081	GTCAGTCACTCCCTGGTCACCTTCGAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTT	1140
Db	8902	GTCGCGAGCTCCCTGGTCACCTTCGAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTT	8961
Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAAACGCCAAAAAGAGGGGAACTCTTTATTTTATA	1200

Db	8962	CAAAATCGAAGAGCTTTAGACTTGCTAAACCGCTGAAGAGGGGGAACCTGTTATTTTAA	9021
Qy	1201	GGAGAAGAACGCTGTTATTATGTTAATCAATCAGAAATGTGTCACTGAGAAAGTTAAAGAA	1260
Db	9022	GGGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAAGAA	9081
Qy	1261	ATTGCGAGATCGAATACAAATGATAGACAGAGGAGCTTCNAAACACCCGACGCTGGGGCCTC	1320
Db	9082	ATTGCGAGATCGAATACAAACGATAGACAGAGGAGCTTCGAAACAACTGGACCTGGGGCCTC	9141
Qy	1321	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAAATATTG	1380
Db	9142	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAAATATTG	9201
Qy	1381	TTACTCTCTTTTGGAGCCCTGATCTTTTAAACCTCCTTGTGTTAAGTTGTCTTCTTCAGAAATT	1440
Db	9202	CTACTCTCTTTTGGAGCCCTGATCTTTTAAACCTCCTTGTGTTAAGTTGTCTTCTTCAGAAATC	9261
Qy	1441	GAGCTGTAAAGCTACAGATGGTCTTACAATGGAACCCCA	1481
Db	9262	GAGCTGTAAAGCTACAAATGGAACCCCAAGATGAGTCCAA	9302
RESULT	13	AC145964	184675 bp DNA linear PRI 29-OCT-2003
LOCUS	AC145964	Pan troglodytes BAC clone RP43-12F2 from 7, complete sequence.	
DEFINITION	AC145964		
ACCESSION	AC145964.3	GI:36016769	
VERSION	AC145964.3		
KEYWORDS	HTG.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		
AUTHORS	Shahid,S., Cotton,M., Bielicki,L. and Meyer,R.		
TITLE	The sequence of Pan troglodytes BAC clone RP43-12F2		
REFERENCE	1 (bases 1 to 184675)		
AUTHORS	Sulston,J.E. and Wilson,R.		
TITLE	Sequencing of Pan troglodytes		
JOURNAL	Unpublished (2001)		
REFERENCE	3 (bases 1 to 184675)		
AUTHORS	Wilson,R.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	4 (bases 1 to 184675)		
AUTHORS	Wilson,R.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-SEP-2003) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	5 (bases 1 to 184675)		
AUTHORS	Wilson,R.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-SEP-2003) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	6 (bases 1 to 184675)		
AUTHORS	Wilson,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-OCT-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA		
COMMENT	On Sep 26, 2003 this sequence version replaced gi:35073486.		
	----- Genome Center		
	Center: Washington University Genome Sequencing Center		
	Center code: WUGSC		
	Web site: http://genome.wustl.edu		
	Contact: submissions@watson.wustl.edu		
	----- Summary Statistics		
	Center project name: C_PT012F02		

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Clint', Verkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen (<http://www.resgen.com>) or Pieter de Jong and co-workers at <http://www.bacpac.chori.org>.

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES		Location/Qualifiers									
source		1. 184675	/organism="Pan troglodytes"								
		/mol_type="genomic DNA"									
		/db_xref="taxon:9598"									
		/chromosome="7"									
		/map="7"									
		/clone="RP43-12P2"									
		/clone_lib="RPCI-43"									
unsure		115419..115500	/note="Sequence derived from one plasmid subclone."								
unsure		138694..138701	/note="Sequence derived from one plasmid subclone."								
misc_feature		140277..140388	/note="Sequence derived from PCR product of project DNA."								
ORIGIN											
Query Match		90.6%;	Score	1341.8;	DB	9;	Length	184675;			
Best Local Similarity		94.1%;	Pred.	No.	0;						
Matches		1394;	Conservative	0;	Mismatches	87;	Indels	0;	Gaps	0;	
Qy	1	ATGGCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTACCCCTTTTCACTCTCACT	60								
Db	153623	ATGGCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCCTTTTCACTCTCACT	153682								
Qy	61	GCACCCCTCCATCTGCTGACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA	120								
Db	153683	GCACCCCTCCATCGCGCTGTATGACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA	153742								
Qy	121	ACGGGGTCTTCGGAAATATTGATGCCCCCATCATATAGGAGTTTATTAAGGGAAATCC	180								
Db	153743	ATGCAGCGTCCCGAAATATTGATGCCCCCATCATATAGGAGTTTCTTAAGGGAAATCC	153802								
Qy	181	ACCTTCACTGCCACACCCATATGCGCGGCAACTGCTATAACTCTGCCACTCTTTGCATG	240								
Db	153803	ACCTTCACTGCCACACCCATATGCGCGGCAACTGCTATCACTCTGCCACTCTTTGCATG	153862								
Qy	241	CATGCAAACTACTATTATTGACAGGGAAATGATTAATCTAGTTTGTCTCGAGGACTT	300								
Db	153863	CATGCAAACTACTATTATTGACAGGGAAATGATTAATCTAGTTTGTCTCGAGGACTT	153922								
Qy	301	GGAGCCACTGTCTGTGGACTTACTTACCCATACCCAGTAGTCTGATGGGGTGGAAAT	360								
Db	153923	GGAGTCACTGTCTGTGGACTTACTTACCCATACCCAGTAGTCTGATGGGGTGGAGTT	153982								

Qy	361	CAAGGTCAGCAAGAGAAAAAAGTAAGAAAGCAATCTCCCAACTGACCCGGGACAT	420
Db	153983	CAAGTCAAGCAAGAGAAAAAAGTAAGAAAGCAATCTCCCAACTGACCCGGGACAT	154042
Qy	421	AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAACTCATATGAAGAACTCCGTTACC	480
Db	154043	AGCACCTTAGGCCCTACAAAGGACTAGATCTCTCAAACTCATATGAAGAACTCCGTTACC	154102
Qy	481	CATCTCGCTGTGAGCCCTATTATACCACTCTCGGCTCCATAGGCTCTCAGCC	540
Db	154103	CATCTCGCTGTGAGCCCTATTATACCACTCTCGGCTCCATAGGCTCTCAGCC	154162
Qy	541	CAAAACCTCTACTACTGTTGGATGTGCTCTCCCTGCACTTCAGGCCATACATTCAATC	600
Db	154163	CAAAACCTCTACTACTGTTGGATGTGCTCTCCCTGCACTTCAGGCCATATGTTCAATC	154222
Qy	601	CCTGTTCTGTAACAAATGGAACTTCAGACAGAAATAAACACCACTTCGGTTTATGTA	660
Db	154223	CCTGTTCTGTAACAAATGGAACTTCAGACAGAAATAAACACCACTTCGGTTTATGTA	154282
Qy	661	GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCAGCTGTGTAATTT	720
Db	154283	GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCAGCTGTGTAATTT	154342
Qy	721	AGCAATCTATAGACACAAACCACTCCCAATGCAATGCAATGCAATGCAATGCAATG	780
Db	154343	AGCAATCTATAGACACAAACCACTCCCAATGCAATGCAATGCAATGCAATGCAATG	154402
Qy	781	ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCTATCATGTTTG	840
Db	154403	ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCTATCATGTTTG	154462
Qy	841	AATGCTCTTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900
Db	154463	AATGCTCTTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	154522
Qy	901	ACTGAAACAAGATTATTAACAATCATGTCTTAAAGCCCCCAACAAAGAGTACCCATT	960
Db	154523	ACTGAAACAAGATTATTAACAATCATGTCTTAAAGCCCCCAACAAAGAGTACCCATT	154582
Qy	961	CTTCTTTTGTATCAGAGCAGGAGTGTAGCAGACTTAGTCTGCGCCCTTATGACCATCTAC	1020
Db	154583	CTTCTTTTGTATCAGAGCAGGAGTGTAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT	154642
Qy	1021	ACAACCTCTACTCAGTCTTCTACTACAACTATCTCAAGAAATAAATGGTACATGGAACAG	1080
Db	154643	ACAACCTCTACTCAGTCTTCTACTACAACTATCTCAAGAAATAAATGGTACATGGAACAG	154702
Qy	1081	GTCACTGACTCCCTGGTCAAGATCAACTTAACTCTCCTAGCAGCAGTAGTCTCTT	1140
Db	154703	GTCCGCGACTCCCTGGTCAAGATCAACTTAACTCTCCTAGCAGCAGTAGTCTCTT	154762
Qy	1141	CAAAATCGAAGAGCTTTAGACTGCTTAACCGCCAAAGAGGGGAACTGTTTATTTTAA	1200
Db	154763	CAAAATCGAAGAGCTTTAGACTGCTTAACCGCCAAAGAGGGGAACTGTTTATTTTAA	154822
Qy	1201	GGAGAAAGACGCTGTTATTTATGTTAATCAATCCAGAAATGTCTCAGAGAAAGTAAAGAA	1260
Db	154823	GGAGAAAGATGCTGTTATTTATGTTAATCAATCCAGAAATGTCTCAGAGAAAGTAAAGAA	154882
Qy	1261	ATTTCGAGATCGAATCAATGTTAGACAGAGAGCTTCAAAACACCGAACCTGGGCGCTC	1320
Db	154883	ATTTCGAGATCGAATCAATGTTAGACAGAGAGCTTCAAAACACCGAACCTGGGCGCTC	154942
Qy	1321	CTCAGCAATGATGCTCTGCT	1380
Db	154943	CTCAGCAATGATGCTCTGCT	155002
Qy	1381	TTACTCTCTTTGGAGCCCTGTATCTTTAACTCTTGTGTTAAGTTTGTCTCTCCAGAAAT	1440
Db	155003	CTACTCTCTTTGGAGCCCTGTATCTTTAACTCTTGTGTTAAGTTTGTCTCTCTCCAGAAAT	155062

QY	1441	GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA	1481	Db	7819	ATGGCCCTCCCTTATCATATATTTTCTCTTACTGTTCTTTTACCCTCTTCTCCTCACT	7878
Db	155063	GAAGCTGTAAAGCTACAAATGGAACCCCAAGATCAGTCCAA	155103	QY	61	GCACCCCTCCCTGCTGTGTACACACAGTAGTCCCTTACCAGAGTTTCTATGAAGA	120
				Db	7879	GCACCCCTCCCTGCTGTGTATGACCACTAGTCCCTTACCAGAGTTTCTATGAAGA	7938
RESULT 14				QY	121	ACGGGCTTCTGGAAATATTGATGCCCCCATATATAGAGTTTATCTAAGGGAATCC	180
AY101586		10229 bp DNA linear PRI 11-FEB-2004		Db	7939	ATGACGCTCCCGGAAATATTGATGCCCCCATATATAGAGTTTCTAAGGGAACCC	7998
DEFINITION		Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVW1		QY	181	ACCTTCACTGCCCCACACACCATATGCCCGCAACTGCTATAACTCTCCCACTCTTTGCATG	240
ACCESSION		AY101586		Db	7999	ACCTTCACTGCCCCACACCATATGCCCGCAACTGCTATACTCTGCACTCTTTGCATG	8058
VERSION		AY101586.1 GI:37544407		QY	241	CATGCAATATCTATTATTGGACAGGGAATATTAATCTTAGTTGTCTCTGAGGACTT	300
KEYWORDS				Db	8059	CATGCAATATCTATTATTGGACAGGGAATATTAATCTTAGTTGTCTCTGAGGACTT	8118
SOURCE		Pan troglodytes (chimpanzee)		QY	301	GGAGCCACTGTCTGTGGACTTACTTTCACCATACAGTATGTCTGATGGGGTGAATTT	360
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		Db	8119	GGAGTCACTGTCTGTGGACTTACTTTCACCATACAGTATGTCTGATGGGGTGAATTT	8178
REFERENCE		Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.		QY	361	CAAGGTCAAGCAAGAAAACAAGTAAAGGAAGCAATCTCCCACTGACCCCGGACAT	420
AUTHORS		The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology		Db	8179	CAAGATCAGGCAAGAGAAAACATGTAAAAGAGTAACTCTCCCACTGACCCCGGTACAT	8238
TITLE		Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		QY	421	AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAAACTACATGAAACCTCCGTAAC	480
JOURNAL		14757826		Db	8239	AGGCCCTTAGGCCCTTACAAAGACTAGATCTCTCAAAACTACATGAAACCTCCGTAAC	8298
PUBMED				QY	481	CATACTCGCTGTGTAGCTATTTAATACACCTCACTCGGCTCCATGAGTCTCAGCC	540
REFERENCE		Mallet, F., Bouton, O. and Oriol, G.		Db	8299	CATACTCGCTGTGTAGCTATTTAATACACCTCACTCGGCTCCATGAGTCTCAGCC	8358
AUTHORS		Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France		QY	541	CAAAACCTTACTTAACCTGTTGGATATGCTCCCTCACTTCAAGGCTATGTTTCAATC	600
TITLE		Location/Qualifiers		Db	8359	CAAAACCTTACTTAACCTGTTGGATATGCTCCCTCACTTCAAGGCTATGTTTCAATC	8418
FEATURES		1. .10229		QY	601	CTGTGTTCTGAAACAATGGAACAACCTTCAAGCAGAGAAATAAACACCACTTCCGTTT	660
source		/organism="Pan troglodytes"		Db	8419	CTGTGTTCTGAAACAATGGAACAACCTTCAAGCAGAGAAATAAACACCACTTCCGTTT	8478
		/mol_type="genomic DNA"		QY	661	GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT	720
		/isolate="1"		Db	8479	GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT	8538
		/isolation_source="PBMC of chimpanzee 1"		QY	721	AGCAATCTATAGACACACACAGCTCCCAATGATCAGGTGGGTACACCTCCACACGA	780
		/db_xref="taxon:9598"		Db	8539	AGCAATCTATAGACACACACAGCTCCCAATGATCAGGTGGGTACACCTCCACACGA	8598
		/cell_type="PBMC"		QY	781	ATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCATTTGTTG	840
		/note="isolated by PCR; ERVW1 locus, allele A"		Db	8599	ATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCATTTGTTG	8658
		endogenous_virus: ERV-W"		QY	841	AATGGCTCTTCAAGATCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900
		1. .779		Db	8659	AATGGCTCTTCAAGATCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	8718
		join(248. .810,2889. .9820)		QY	901	ACTGAACAAGATTTATACATCATGTCTGTTTAAAGCCCAACCAACCAAGAGTACCCATT	960
LTR		/note="putative mRNA transcript 1; gag-pol pseudogene"		Db	8719	ACTGAACAAGATTTATACATCATGTCTGTTTAAAGCCCAACCAACCAAGAGTACCCATT	8778
mRNA		join(248. .810,7592. .9820)		QY	961	CTTCTCTTTCTTATCAGACGAGGTGCTAGGACAGCTAGGTACTGTCATTTGGCAGTATC	1020
mRNA		/product="envelope glycoprotein"		Db	8779	CTTCTCTTTCTTATCAGACGAGGTGCTAGGTGCTAGGTACTGTCATTTGGCAGTATC	8838
mRNA		/note="putative mRNA transcript 2"		QY	1021	ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGCATGGAAACAG	1080
CDS		join(248. .810,9255. .9820)		Db	8839	ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGCATGGAAACAG	8898
		/note="putative mRNA transcript 3; no coding region"		QY	1081	GTCACTGACCTCCCTGGTCCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTAGTCTT	1140
		7819. .9435		Db	8899	GTCCGCGACTCCCTGGTCCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTAGTCTT	8958
		/note="syncytin"					
		/codon_start=1					
		/product="envelope glycoprotein"					
		/protein_id="AA68165.1"					
		/db_xref="GI:37544408"					
		/translat="MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQBFLLRMORP					
		GNIDAPSRFSKPTPTFTAHMPKNCYHSATLCHNANTHYWTGKINPSCPGILGV					
		TVCWYFTHGMSDGGVQDQAREKHVKEVISQLTRVHSARPKGLDLSKLHETRT					
		HTRLVSLFNTLTGLHEVSAQNPNWCILPFRPVSIPVPEQWNNFSTEINTVTS					
		LVGLVNLNLTHTSNLTCKVFNNTDNTNSQIRWVTPPTQIVCLPSGIFVFCGTS					
		YRLNGSSMCFSLFVPPMTIYTEQDLYNYVSKPRKRVFIPFVIGAGVILG					
		TGIGITTSFTQFYKLSQELNGDMERVADSLVTLQDOLSLAAVVLQNRALLDLTAE					
		RGCTCLPLGECYVYNGSIGVTEKVKKEIRDIQRRAEELRNTGPMGLSQWMPWILP					
		FLGPLAIIILLFGLPCIFNLLVNFVSSRIEAVKLMQEPKMQSKTKLYRPLDRPASP					
		RSVDNDIKGTPPPEILLTAQPLLRNSAGSS"					
		9494. .10229					
ORIGIN							
Query Match		90.5%; Score 1340.2; DB 9; Length 10229;					
Best Local Similarity		94.1%; Pred. No. 0;					
Matches 1393; Conservative 0; Mismatches 88; Indels 0; Gaps 0;							
QY	1	ATGGCCCTCCCTTATCATATATTTTCTCTTACTGTTCTTTTACCCTCTTCTCCTCACT	60				

QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACGCGCAAGAGGGGAACTGTTTATTTTA 1200
Db 8959 CAAATCGAAGAGCTTTAGACTTGTCTAAACGCGCGAAGAGGGGAACTGTTTATTTTA 9018
QY 1201 GGAGAGAAACGCTGTTTATTTATTTATTAATCAATCCAGAAATGTCACATGAGAAAGTTAAAGAA 1260
Db 9019 GCGGAGAAATGCTGTTTATTTATTTATTAATCAATCCGAAATCGTCTGAGAAAGTTAAAGAA 9078
QY 1261 ATTCGAGATCGAATCAATGTAAGACAGAGAGCTTCAAAACACCGAAGCTGCGGCCTC 1320
Db 9079 ATTCGAGATCGAATCAATGTAAGACAGAGAGCTTCAAAACACCGAAGCTGCGGCCTC 9138
QY 1321 CTCAGCAATGGATGCGCTGGGTTCTCCCTTTCTTAGACCTCTAGCAGCTCTAATATG 1380
Db 9139 CTCAGCAATGGATGCGCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATG 9198
QY 1381 TTACTCTCTTTGGACCTGATCTTTAACTCTCTGTTAACTTTGTTCTCTTCCAGAATT 1440
Db 9199 CTACTCTCTTTGGACCTGATCTTTAACTCTCTGTTAACTTTGTTCTCTTCCAGAATC 9258
QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 9259 GAAGCTGTAAGCTACAAATGAGCCCAAGATGCAGTCCAA 9299

RESULT 15
AY101587 10229 bp DNA linear PRI 11-FEB-2004
LOCUS Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVW1
DEFINITION locus, allele B, complete sequence.
ACCESSION AY101587
VERSION AY101587.1 GI:37544409
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10229)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 10229)
Mallet,F., Bouton,O. and Oriol,G.
Direct Submission
AUTHORS Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-BioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
JOURNAL Location/Qualifiers
FEATURES
source
1..10229
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/isolate="1"
/isolation_source="PBMC of chimpanzee 1"
/db_xref="taxon:9598"
/cell_type="PBMC"
/note="isolated by PCR; ERVW1 locus, allele B
endogenous_virus: ERV-W"
1..779
join(248..810,2889..9820)
/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..810,7592..9820)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..810,9255..9820)
/note="putative mRNA transcript 3; no coding region"
7819..9435
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"

/protein_id="AAM68166.1"
/db_xref="GI:37544410"
/translation="MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYQBFLLWRMORP
GNIDAPSYRSFKSGTFTTAHTHMPNCHVSATLCKHANTHRYPKIGLIDLSKHELTRT
TVCTWYFTHTGMSDGGVQDQAREKHVKWICLPLNFRPYVSIPLVPEQWNPFSEINTSV
HTLVLSEINTLTGLHEVSAQNTNCWICLPLNFRPYVSIPLVPEQWNPFSEINTSV
LVGLVSNLEIHTHTSNLTCVKFSNTTDDTNSQCIRWVTPPTQVPSGFFVCGTSA
YRCINGSSSEMSCLFLVPMWTIYTBODLYNVVYSPKRNKRVPLIFVIGAGVLGALG
TGIGGTTSTQFYKLSQELNGDMERVADSLVLODLNSLAALVQNRRLDILLTAE
RGCTCLFLGEECCYVNGSGIVTEKVEIRDRLOBAEELRNTGPNGLLSQWMPWLP
FLGLAAIILLLLFGPCIFNLVNFVSSRIEAVLQNEPKMQSKTKIYRPLDRPASP
RSDVNDIKGTPPEIILTAPQLLRPNAGS"
9494..10229
LTR
ORIGIN
Query Match 90.5%; Score 1340.2; DB 9; Length 10229;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTATCATACTTTTCTTTACTGTCTCTTACCCCTTTGGCTCTCACT 60
Db 7819 ATGGCCCTCCCTATCATATTTTCTTTACTGTCTTTTACCCTCTTTCACCTCACT 7878
QY 61 GCACCCCTCCATGCTGTFACACCAAGTAGTCCCTTACCAAGAGTTTCTATGAGA 120
Db 7879 GCACCCCTCCATGCTGTFATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGAGA 7938
QY 121 ACGGGCTTCTGGAATATTTGATGCCCCCATATATAGGAGTTTATCTAAGGGGAACTCC 180
Db 7939 ATGCAGCGTCCCGGAATATTTGATGCCCCCATATATAGGAGTTTCTAAGGGGAACTCC 7998
QY 181 ACCTTCACTGCGCCACACCCATATGCCCCGCAACTGTATAACTCTGCACTCTTTGCAATG 240
Db 7999 ACCTTCACTGCGCCACACCCATATGCCCCGCAACTGTATCACTCTGCACTCTTTGCAATG 8058
QY 241 CATGCAATATCTCATTTATTTGACAGGGAATATGATTAATCTAGTTCTCTCGGGAGGACTT 300
Db 8059 CATGCAATATCTCATTTATTTGACAGGGAATATTAATCTAGTTCTCTCGGGAGGACTT 8118
QY 301 GGAGCCTGCTGTTGGACTTCTTACCCCATACCACTACAGTATGCTGTATGTTGGGGTGGAAAT 360
Db 8119 GGAGTCACTGCTGTTGGACTTCTTACCCCATACTGTTATGCTGTATGTTGGGGTGGAGTT 8178
QY 361 CAAAGTCAGGCAAGAGAAAAAACAAGTAAAGAAAGCAATCTCCCACTGACCCGGGGACAT 420
Db 8179 CAAAGTCAGGCAAGAGAAAAAACAAGTAAAGAAAGCAATCTCCCACTGACCCGGGGTACAT 8238
QY 421 AGCACCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCTCCGTACC 480
Db 8239 AGCGCCTTAGGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACC 8298
QY 481 CATACTCGCTGTTGAGCTATTTAATACCACTCTCACTCGGCTCCATGAGTCTTCAGCC 540
Db 8299 CATACTCGCTGTTGAGCTATTTAATACCACTCTCACTCGGCTCCATGAGTCTTCAGCC 8358
QY 541 CAAAACCTTACTAATGTTGGATGCTCCCTCCCTGCACTTCAGGCCATACATTTCAATC 600
Db 8359 CAAAACCTTACTAATGTTGGATATGCTCCCTCCCTGAACTTTCAGGCCATATGTTTCAATC 8418
QY 601 CTGTTCTGACCAATGGAACAATCTCAGCAGAGAAATAAACCACCTTCCGTTTTAGTA 660
Db 8419 CTGTTCTGACCAATGGAACAATCTCAGCAGAGAAATAAACCACCTTCCGTTTTAGTA 8478
QY 661 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 720
Db 8479 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 8538
QY 721 AGCAATATATAGACACCAACAGCTCCCAATGCAATCAGTGGTGAACCTCCACACCA 780
Db 8539 AGCAATATATAGACACCAACAGCTCCCAATGCAATCAGTGGTGAACCTCCACACCA 8598
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTGTTGTTACCTCAGCCTATCATTTGTTG 840

8599	ATAGTGTGCTACCTCAGGAATATTTTGTGTGTGTACCTCAGCTATCGTTGTTG	8658	
841	AATGGCTCTTCAGAACTATATGCTTCTCTCTCATTTAGTGCCCTATGACCATCTAC	900	
8659	AATGGCTCTTCAGAACTATATGCTTCTCTCTCATTTAGTGCCCTATGACCATCTAC	8718	
901	ACTGAACAAGATTTATACAATCATGCTGTAAGGCCCAACAACAAAGAGTACCCATT	960	
8719	ACTGAACAAGATTTATACAATCATGCTGTAAGGCCCAACAACAAAGAGTACCCATT	8778	
961	CTTCTCTTTTATCAGACGAGGAGTGTAGGACAGCTAGTACTGTCATTTGGCAGTATC	1020	
8779	CTTCTCTTTTATCAGACGAGGAGTGTAGGACAGCTAGTACTGTCATTTGGCAGTATC	8838	
1021	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG	1080	
8839	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG	8898	
1081	GTCACTAGTCCCTGGTCAACCTTGAAGATCAACTTAACTCCCTAGCAGAGTATGCTTT	1140	
8899	GTCCGCGACTCCCTGGTCAACCTTGAAGATCAACTTAACTCCCTAGCAGAGTATGCTTT	8958	
1141	CAAAATCGAAGAGCTTTAGACTTGTACCCGCAAGAGAGGGGAACTGTTATTTT	1200	
8959	CAAAATCGAAGAGCTTTAGACTTGTACCCGCAAGAGAGGGGAACTGTTATTTT	9018	
1201	GGAGAAGACGCTGTTATTTATGTTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA	1260	
9019	GGGAAGAATGCTGTTATTTATGTTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA	9078	
1261	ATTGAGATCGAATCAATGTTAGACGAGAGGAGTCTCAAAACACCGAACGCTGGGGCTC	1320	
9079	ATTGAGATCGAATCAATGTTAGACGAGAGGAGTCTCAAAACACCGAACGCTGGGGCTC	9138	
1321	CTGAGCAATGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	1380	
9139	CTGAGCAATGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	9198	
1381	TTACTCTCTTTGGACCCGTATCTTTAACTCTCTTTAAAGTTTGTCTCTCCAGAAAT	1440	
9199	CTACTCTCTTTGGACCCGTATCTTTAACTCTCTTTAAAGTTTGTCTCTCCAGAAAT	9258	
1441	GAAGCTGTAAGCTCAGATGCTTCTTCAAAATGGAACCCCA	1481	
9259	GAAGCTGTAAGCTCAGATGCTTCTTCAAAATGGAACCCCA	9299	
RESULT 16			
AY101588	10230 bp	DNA	linear
LOCUS	Gorilla gorilla isolate 1	endogenous retrovirus ERV-W, ERVW1	
DEFINITION	locus, allele A, complete sequence.		
ACCESSION	AY101588		
VERSION	AY101588.1	GI:37544411	
KEYWORDS	Gorilla gorilla (gorilla)		
SOURCE	Gorilla gorilla		
ORGANISM	Gorilla gorilla		
REFERENCE	Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.		
AUTHORS	1 (bases 1 to 10230) Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.		
TITLE	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
PUBLISHED	14757826		
REFERENCE	2 (bases 1 to 10230) Mallet, F., Bouton, O. and Oriol, G.		
AUTHORS	Direct Submission		
TITLE	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142		
JOURNAL	CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée		
REFERENCE	d'Italie, Lyon 69364 cedex 07, France		
FEATURES	Location/Qualifiers		


```
QY 181 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
Db 7986 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 8045
QY 241 CATGCAAAATCTCATTTATGGACAGGAAATGATTAATCTCTAGTTGTCTGAGGACTTT 300
Db 8046 CATGCAAAATCTCAGTATTGGACAGGAAATGATTAATCTCTAGTTGTCTGAGGACTTT 8105
QY 301 GGAGCCACTGTCTGTTGGACTTACTTACCCATACAGTATGCTCATGGGGGTGGAATT 360
Db 8106 GGAGTCACTGTCTGTTGGACTTACTTACCCATACAGTATGCTCATGGGGGTGGAATT 8165
QY 361 CAAAGTTCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 8166 CAAAGTTCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8225
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
Db 8226 AGCACCTTAGCCCTTACAAAGGACTAGTCTCTCAAACTACATGAAACCCCTCCGTACC 8285
QY 481 CATACTCGCTGTGAGCTTATTAATACACCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 8286 CACACTCGCTGTGAGCTTATTAATACACACTCACTCGGCTCCATGAGGTCTCGGCC 8345
QY 541 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACTTCAGGGCCATACATTTCAATC 600
Db 8346 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACTTCAGGGCCATATGTTCAATC 8405
QY 601 CTTGTTCTGCAACAAATGGAACAACTTCAGCACAGAAATAAACAACCACTTCCTGTTTAGTA 660
Db 8406 CTTGTTCTGCAACAAATGGAACAACTTCAGCACAGAAATAAACAACCACTTCCTGTTTAGTA 8465
QY 661 GGACCTCTGTTTCCAACTCTGGAATAACCCATCTCAACCTCACTGCTGTAAATTT 720
Db 8466 GGACCTCTGTTTCCAACTCTGGAATAACCCATCTCAACCTCACTGCTGTAAATTT 8525
QY 721 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACAG 780
Db 8526 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAACTCTCCACACAG 8585
QY 781 ATAGTCTGCTACCCCTCAGGAATATTTTTGTCTGTTGTTACCTCAGCCTATCATTTGTTG 840
Db 8586 ATATTCTGCTACCCCTCAGGAATATTTTTGTCTGTTGTTACCTCAGCCTATCATTTGTTG 8645
QY 841 AATGGCTCTCAGAAATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 8646 AATGGCTCTCAGAAATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8705
QY 901 ACTGAAACAAGATTTATCAATCATGCTGTAACCTAAGCCCAACCAACAAAGAGTACCCATT 960
Db 8706 ACTGAAACAAGATTTATCAATCATGCTGTAACCTAAGCCCAACCAACAAAGAGTACCCATT 8765
QY 961 CTTCTCTTTTGTATCAGACAGGAGTCTAGGCAGACTAGGTACTGCGCATGCGCAGTATC 1020
Db 8766 CTTCTCTTTTGTATGAGGAGCAGGAGTCTAGGTGCTAGGTACTGCGCATGCGCAGTATC 8825
QY 1021 ACAACCTCTACTGTTCTACTACAACTATCTCAAGAAATAATGTTGATGATGAGACAG 1080
Db 8826 ACAACCTCTACTGTTCTACTACAACTATCTCAAGAAATAATGTTGATGATGAGACAG 8885
QY 1081 GTCACTGACTCCCTGGTCACTGTCGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCTTT 1140
Db 8886 GTCCCGGACTCCCTGGTCACTGTCGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCTTT 8945
QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAAGAGGGGAAACCTGTTATTTTAA 1200
Db 8946 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAAACCTGTTATTTTAA 9005
QY 1201 GGAGAGAACGCTGTTATTTATGTTAAATCAATCCAGAAATGTCACCTAGAGAAAGTTAAGAA 1260
Db 9006 GGGAGAGAAAGCTGTTATTTATGTTAAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAA 9065
QY 1261 ATTCGAGATCGAATACATGTTAGAGCAGAGGAGCTTCAAAACACCGAACGCTGGGGCTC 1320
```

```
Db 9066 ATTGAGATCGAATAACAACGTAGACGAGGAGCTTCGAAACATTTGGACCCCTGGGCGCTC 9125
QY 1321 CTCACCCAAATGATCCCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 9126 CTCACCCAAATGATCCCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTATATATTG 9185
QY 1381 TTACTCTCTCTTTGGACCCCTGTATCTTTAACTCTCTGTTAAAGTTGTTCTCTCCAGAA 1440
Db 9186 CTGCTCTCTTTGGACCCCTGTATCTTTAACTCTCTGTTAACTCTCTCTCTCTCCAGAA 9245
QY 1441 GAAGCTGTAAAGCTACAGATGCTTACAAATGGAACCCCA 1481
Db 9246 GAAGCTATAAACTACAAATGAGCCCAAGATGGAGTCCCA 9286

RESULT 18
AY101592 10248 bp DNA linear PRI 11-FEB-2004
LOCUS Hylobates pileatus isolate 1 endogenous retrovirus ERV-W, ERVWE1
DEFINITION locus, allele A, complete sequence.
ACCESSION AY101592
VERSION AY101592.1 GI:37544419
KEYWORDS Hylobates pileatus (pileated gibbon)
ORGANISM Hylobates pileatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE 1 (bases 1 to 10248)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCES 2 (bases 1 to 10248)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
DIRECT SUBMISSION
TITLE Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
JOURNAL CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..10248
/organism="Hylobates pileatus"
/mol_type="genomic DNA"
/isolate="1"
/isolation_source="PBMC of gibbon 1"
/db_xref="taxon:9589"
/cell_type="PBMC"
/notes="isolated by PCR; ERVWE1 locus, allele A
endogenous_virus: ERV-W"
1..773
join(248..804,2997..9809)
/notes="putative mRNA transcript 1; gag-pol pseudogene"
join(248..804,7581..9809)
/product="envelope glycoprotein"
/notes="putative mRNA transcript 2"
join(248..804,9244..9809)
/notes="putative mRNA transcript 3; no coding region"
7808..9424
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68171.1"
/db_xref="GI:37544420"
/translation="MALPFIHFLFTVLLPSFTLTAPPCCRTSSSPYQEFPLARTORP
GNIDAPLYRSFKSGSTFTATYMPRTYNSATLCMHANTQYMTGMINPSCGGLGV
TVCWYFTHTGMSDGGVDQARKEHKEVISQLTQVHSTSPYKGLDLSKLHETLRT
HRLVSLFNLTTLGLHEVAQNPTNCWMLPLDFRPYVSTPVEQWNFTINTTSV
LVGPLVSLNLEITHNTLCTVKFSNTDNTNSQICIRWTPTRIFCLPSGTFPVCSTSA
YRCINGSSSMCFSLFLVPWMTIYEQDLNYYVVKPRKRVPILPVWAGVLGALG
TGCSTGTTSTQFYIKLSRELNGDMERVADSLVTLQDQLSLAAVLPQNRALLDITAE
RGCTCLFLGECCYYNQSGIVTEKVKIRDRIQRRABEELRNIGPWLGLSQQWPFWILP
```

LTR ORIGIN		FLGPLAAIILLLLFGPCIFINLLVNFVSSRIEAIKLQMPKMSKTKNVRSLDWPASP RSDVNDIKIGIPPEEISTAQPLLRPNISAGSS" 9483..10248	
Query Match		90.4%; Score 1338.6; DB 9; Length 10248;	
Best Local Similarity		94.0%; Pred. No. 0;	
Matches 1392; Conservative		0; Mismatches 89; Indels 0; Gaps 0;	
Qy	1	ATGGCCCTCCCTATCATACTTTTCTTTACTGTCTCTTACCCCTTTGGCTCTCACT	60
Db	7808	ATGGCCCTCCCTATCATATTTTCTTTACTGTCTCTTACCCCTTTGGCTCTCACT	7867
Qy	61	GCACCCCTCCATGTCTGTACACCCAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA	120
Db	7868	GCACCCCTCCATGCCGCTGTATGACCACTAGCTCCCTTACCAAGAGTTTCTATGGAGA	7927
Qy	121	AGCGGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC	180
Db	7928	AGCAGCGTCTGGAAATATTGATGCCCATTTGTATAGAGTTTCTAAGGGAAGCCCC	7987
Qy	181	ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG	240
Db	7988	ACCTTCACCTGCCACACCTATATGCCCGCACTGCTATAACTCTGCCACTCTTTGCAATG	8047
Qy	241	CATGCAAACTACTCATTTATGGACAGGGAATAATTAATCTTAGTTGCTCGGAGACTT	300
Db	8048	CATGCAAACTACTCAGTATTGGACAGGGAATAATTAATCTTAGTTGCTCGGAGACTT	8107
Qy	301	GGAGCCACTGCTGTGGACTTACTTCAACCATACCAGTATGCTGATGGGGTGGAAAT	360
Db	8108	GGAGTCACTGCTGTGGACTTACTTCAACCATACCAGTATGCTGATGGGGTGGAGTT	8167
Qy	361	CAAGGTCAGGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCAACTGACCCGGGACAT	420
Db	8168	CAAGATCAGCAAGAGAAAAACATGTAAGAAGTAATCTCCCAACTGACCCAGGTACAT	8227
Qy	421	AGCACCCCTAGCCCTACAAAGACTAGTTCTCTCAAACTACATCAAAACCTCCCTACC	480
Db	8228	AGCACCTTAGCCCTACAAAGACTAGATCTCTCAAACTACATCAAAACCTCCCTACC	8287
Qy	481	CATACTCGCTGGTGACCTATTATAACCAACCTCACTCGGCTCCATGAGGTCTCAGCC	540
Db	8288	CACACTCGCTGGTAAGCCATTATTAACACACTCACTGGGCTCCATGAGGTCTCGGCC	8347
Qy	541	CAAAACCTACTAACTGTTGGATGCTCCCTCGCACTTCAGGGCATACATTTCAATC	600
Db	8348	CAAAACCTACTAACTGTTGGATGCTCCCTCGCACTTCAGGGCATATGTTTCAATC	8407
Qy	601	CTGTCTTGNACAAATGGAACTTTCAGCACAGAAATAAACACCACTTCGGTTTAGTA	660
Db	8408	CTGTCTTGNACAAATGGAACTTTCAGCACAGAAATAATAACCACTTCGGTTTAGTA	8467
Qy	661	GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCACTGTGTAATAAT	720
Db	8468	GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCACTGTGTAATAAT	8527
Qy	721	AGCAATACTATAGACAAACAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGA	780
Db	8528	AGCAATACTATAGACAAACAGCTCCCAATGCATCAGGTGGGTAACTCTCCACACGA	8587
Qy	781	ATAGTCTGCCTACCTCAGGAATATTTTGTGCTGTGTACCTCAGGCTATCATTTGTTG	840
Db	8588	ATATTCGCCCTACCTCAGGAATATTTTGTGCTGTGTACCTCAGGCTATCGTTGTTG	8647
Qy	841	AATGGCTCTTCAGAACTATGCTCTCTCATCTTAGTGGCCCCCTATGACCATCTAC	900
Db	8648	AATGGCTCTTCAGAACTATGCTCTCTCATCTTAGTGGCCCCCTATGACCATCTAC	8707
Qy	901	ACTGACAGAGTTTATACAAATCATGTGCTAGCTTAAGCCCCCAACAAGAGTACCCATT	960
Db	8708	ACTGACAGAGTTTATACAAATCATGTGCTAGCTTAAGCCCCCAACAAGAGTACCCATT	8767
Qy	961	CTTCCCTTTTCTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGGCATTTGCAGTATC	1020
Db	8768	CTTCCCTTTTCTATGGAGCAGGAGTCTAGGTGCACTAGGTACTGGCATTTGCAGTATC	8827
Qy	1021	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGACAG	1080
Db	8828	ACAACTCTACTCAGTTCTACTACAACTATCTCGAAGAAATAAATGGTGACATGGACAG	8887
Qy	1081	GTCACTGACCTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTT	1140
Db	8888	GTGCGCACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTT	8947
Qy	1141	CAAAATCGAAGAGCTTTAGACTTTGCTAAACCGCAAGAGGGGGAACCTGTTATTTTA	1200
Db	8948	CAAAATCGAAGAGCTTTAGACTTTGCTAAACCGCTGAAAGAGGGGGAACCTGTTATTTTA	9007
Qy	1201	GGAGAAGAAACGCTGTTTATTATGTTAATCAATCCAGATTTGCTCACTGAGAAGTTAAAGAA	1260
Db	9008	GGGGAAGAAATGCTGTTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAAGAA	9067
Qy	1261	ATTGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCGAACGCTGGGGCTC	1320
Db	9068	ATTGAGATCGAATACAAACGTAAGCAGAGAGCTTCAAAACATTTGGACCTGGGGCTC	9127
Qy	1321	CTCAGCCAAATGGATGCGCTCGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATG	1380
Db	9128	CTCAGCCAAATGGATGCGCTCGGATTCCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATG	9187
Qy	1381	TTACTCTCTTTGGACCTCTGATCTTTAAACCTCTGTTAAGTTTGTCTCTCCAGAAAT	1440
Db	9188	CTGCTCTCTTTGGACCTCTGATCTTTAAACCTCTGTTAAGTTTGTCTCTCCAGAAAT	9247
Qy	1441	GAACTGTAAAGCTACAGATGCTCTTCAAAATGGAACCCCA	1481
Db	9248	GAACTGTAAAGCTACAAATGGAACCCCAAGATGAGTCCAA	9288
RESULT 19		AF520477S2 2694 bp DNA linear PRI 11-FEB-2004	
LOCUS		Homo sapiens individual 132 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.	
DEFINITION		AF520478.1 GI:33410924	
ACCESSION		2 of 2	
VERSION		Homo sapiens (human)	
KEYWORDS		Homo sapiens	
SEGMENT		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
SOURCE		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
ORGANISM		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G.,	
AUTHORS		Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.	
TITLE		The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
PUBMED		14757826	
REFERENCE		2 (bases 1 to 2694)	
AUTHORS		Mallet, F., Bouton, O. and Oriol, G.	
TITLE		Direct Submission	
JOURNAL		Submitted (07-JUN-2002) Retrovirology Department, UMR 2142	
FEATURES		CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France	
SOURCE		Location/Qualifiers	
		1..2694	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/isolate="individual 132 allele A"	
		/db_xref="taxon:9606"	
		/chromosome="7"	
		/map="7q21-q22"	
		/sex="female"	
		/cell_type="PBMC's"	

misc_feature	/note="Caucasian										
	endogenous_virus: HERV-W"										
	43..57										
	/note="splice acceptor site"										
	284..1900										
	/note="syncytin"										
	/codon_start=1										
	/product="envelope glycoprotein"										
	/protein_id="AAQ17561.1"										
	/db_xref="GI:33410925"										
misc_feature	/translation="MALPHYHFLTVLLPSFTLTAPPCCRMSTSSPYQEFLEWRMORP GNIDAPYSRLSGKSTFTTAHTRPNCYHSLATCMHANTHYWTGMINPSPCGGLGV TWCVTFQTGMSGDGGVODQAREKHVKEVISOLTRVHGTSYKGLDLSKLEHLRT HTRLVSLNFTLTLGLHEVSAQNPTNCWICLPLNRPVYSIPVPEQNNRSTETNTTSV LVGPLVSNLEIHTSNLTCVKFSNTTYTNSQCIRWTPPTQIVCLPSCGIFVFCGTSIA YKLNSSSMSCPLSPVPMWTITYTEQDLYNYSIKPRNKRVPILFVIGAGVLGALG YRGLGGTSTQFYKLSOBLNGDMERVADSLVTLQDLNSLAALVNLQNRALLDLTAE RGGTCLFLGEBECYVQNGSIYTEKVEIKRDIRIQRARELRNTGPMGLLSOWMPWLPILP FLGPLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQWEPKMQSKTKIYRPLDRPASP RSDVNDIKGTPPEEISAAQPLLRPNAGSS"										
	1705..1720										
	/note="splice acceptor site"										
	1959..2694										
	LTR	ORIGIN									
		Query Match 90.2%; Score 1335.4; DB 9; Length 2694;									
		Best Local Similarity 93.9%; Pred. No. 0;									
		Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;									
		QY	1	ATGCGCCCTCCCTTATCATATCTTTCTCTTACTGTTCTTCTTACCCCTTTTCGCTCTCACT	60						
		DB	284	ATGCGCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCCTTTTCACCTCTCACT	343						
QY		61	GCACCCCTCCCATGCTGCTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA	120							
DB		344	GCACCCCTCCCATGCGCGCTGTATGACCAGTAGCTCCCTTTACCAAGAGTTTCTATGAGA	403							
QY		121	ACGGGGCTTCTTGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC	180							
DB		404	ATGCAGCGTCCCGGAATATTGATGCCCATCTGTATAGGAGTCTTTCTAAGGGAACCCCC	463							
QY	181	ACCTTCACTGCCACACCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTTGCATG	240								
DB	464	ACCTTCACTGCCACACCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATG	523								
QY	241	CATGCAAAATCTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT	300								
DB	524	CATGCAAAATCTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT	583								
QY	301	GGAGCCACTGCTGTGGACTTACTTCAACCATACCATGATGCTGATGGGGTGGAAAT	360								
DB	584	GGAGTCACTGCTGTGGACTTACTTCAACCAAACTGGTATGCTGATGGGGTGGAGTT	643								
QY	361	CAAGGTGAGGCAAGAGAAAAACAAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT	420								
DB	644	CAAGATCAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCCGGGTACAT	703								
QY	421	AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAATPACATGAACACCTCCGTACC	480								
DB	704	GGCACCTTAGCCCTCAAAAGGACTAGATCTCTCAAAATPACATGAACACCTCCGTACC	763								
QY	481	CATAGTCCCTGGTGGAGCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC	540								
DB	764	CATAGTCCCTGGTGAAGCTATTTAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCC	823								
QY	541	CAAAACCTTACTAATGTTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC	600								
DB	824	CAAAACCTTACTAATGTTGGATGTGCTCCCTCGCACTTCAGGCCATATGTTTCAATC	883								
QY	601	CCTGTTCTTGAAACAATGGAACAACTTACAGCAGAAATAAACACCACTTCGTTTATGTA	660								
DB	884	CCTGTACTGAACAATGGAACAACTTACAGCAGAAATAAACACCACTTCGTTTATGTA	943								

QY	661	GGACCTCTTGTTCCTCAATCTCGAATAAACCCTAATACCTCAAACTCCTGCTGTGTAATTT	720							
DB	944	GGACCTCTTGTTCCTCAATCTCGAATAAACCCTAATACCTCAAACTCCTGCTGTGTAATTT	1003							
QY	721	AGCAATATCTATAGACACACACAGCTCCCAATCATCATGAGTGGGTACACCTCCACACAGA	780							
DB	1004	AGCAATATCTATAGACACACACAGCTCCCAATCATCATGAGTGGGTAACTCTCTCCACACA	1063							
QY	781	ATAGTCTGCTTACCTCTCAGGAATATTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	840							
DB	1064	ATAGTCTGCTTACCTCTCAGGAATATTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1123							
QY	841	AATGGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900							
DB	1124	AATGGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1183							
QY	901	ACTGAACAGATTTATACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	960							
DB	1184	ACTGAACAGATTTATACATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1243							
QY	961	CTTCTCTTTTGTATCAGACGAGGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC	1020							
DB	1244	CTTCTCTTTTGTATCAGACGAGGTGCTAGGCAGACTAGGTACTGGCATTGGCAGTATC	1303							
QY	1021	ACAACTCTACTCAGTTCCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAACAG	1080							
DB	1304	ACAACTCTACTCAGTTCCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAACAG	1363							
QY	1081	GTCACTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140							
DB	1364	GTGCGCGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1423							
QY	1141	CAAAATCGAAGACTTTAGACTTGTAAACCGCCAAAGAGGGGGAACCTGTTTATTTTAA	1200							
DB	1424	CAAAATCGAAGACTTTAGACTTGTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTAA	1483							
QY	1201	GGAGAAAGACGCTGTTTATTATGTTAATCAATCAAGAAATTTGCTACTGAGAAAGTTAAAGAA	1260							
DB	1484	GGGGAAGATGCTGTTTATTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGAA	1543							
QY	1261	ATTGAGATCGAATACATGATAGACAGAGGCTTCMAAACACCGAACCTGGGGCTC	1320							
DB	1544	ATTGAGATCGAATACATGATAGACAGAGGCTTCGAAACACTGGAACCTGGGGCTC	1603							
QY	1321	CTCAGCAATGAGTGCCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAATATTG	1380							
DB	1604	CTCAGCAATGAGTGCCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAATATTG	1663							
QY	1381	TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT	1440							
DB	1664	CTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT	1723							
QY	1441	GAAGCTGTAAAGCTACAGATGTTCTTACAAATGGAAACCCCA 1481								
DB	1724	GAAGCTGTAAAGCTACAGATGTTCTTACAAATGGAAACCCCA 1764								

RESULT 20
AF520487S2
LOCUS
DEFINITION
Homo sapiens individual 148 allele B, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
AF520488
AF520488.1 GI:33410944
2 of 2
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2694)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.

AF520487S2 2694 bp DNA linear PRI 11-FEB-2004
Homo sapiens individual 148 allele B, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
AF520488
AF520488.1 GI:33410944
2 of 2
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2694)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.

Db 1424 CAAATCGAAGAGCTTTAGACTTGCTACCGCTGAAGAGGGGAACCTGTTATTTTA 1483
QY 1201 GGAGAGAACGCTGTTATTTATGTTAATCAATCCAGAAATGTCACAGAAAGTTAAAGAA 1260
Db 1484 GGGGAAGAATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1543
QY 1261 ATTGAGATCGAATCAATGTTAGAGCAGAGGAGCTTCAAAACACCGAAGCTGGGGCCTC 1320
Db 1544 ATTGAGATCGAATCAACGTTAGAGCAGAGGAGCTTCGAAACACATGGACCCCTGGGGCCTC 1603
QY 1321 CTCAGCAATGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTTG 1380
Db 1604 CTCAGCAATGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTTG 1663
QY 1381 TTACTCTCTTTGGACCCCTGATCTTAACTCTTGTAAAGTTGTTCTCTTCAGAAATT 1440
Db 1664 CTACTCTCTTTGGACCCCTGATCTTAACTCTTGTAAAGTTGTTCTCTTCAGAAATC 1723
QY 1441 GAAGCTGTTAAGCTACAGATGCTTTACAAATGGAAACCCCA 1481
Db 1724 GAAGCTGTTAAACTCAAAATGGAGCCCAAGATGCAATGCCAA 1764

RESULT 22
AF52050582 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 37 allele A, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520506
VERSION AF520506.1 GI:33410980
KEYWORDS
SEGMENT
SOURCE 2 of 2
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France

FEATURES
source
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 37 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/notes="endogenous_virus: HERV-W"
misc_feature 43. .57
/notes="splice acceptor site"
CDS 284. .1900
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA017575.1"
/db_xref="GI:33410981"
/translation="MALPHYFLFTVLPSFTLTAPPCCRCMTSSSPYOFELWRMORP
GNIDAFSYLSLKGTPFTTAHMPNCHVSATLCMHANHYWTGKMNPSCPGGLGV
TVCWTFYTGMSDGGVQDQAREKHVKEVISQITRVHGTSPPYKGLDLSKHELTUR
HTRLVSLFNTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPEQWNNFSTEINTTSV

LVGLVSNLEIHTHTSNLTCVKFSNTTNTSQICIRWVTPPTQIVCLPSGIFVFCGTS
RYCLNGSSEMCFLSFLVPMWITIEQDLNYNISKPRNKRVLPLFPVIGALVGLG
TGIGITTSQFYKLSQELNGDMVERVADSLVTLDQNLAAVVLQNRALDLITAE
RGCTCLFLGEBECYVUNSGIVTEKVEIRDRIOORABEELRNTGPGWLLSOWMPILP
FLGPDAAIIILLLFGPCIFNLLNFNVSRSRIEAVLQMEPQMOKTKIYRRPLDRPAS
RSDVNDIKGTPPEISAAQPLLRNPSAGSS"
1705. .1720
/note="splice acceptor site"
1959. .2694
misc_feature
LTR
ORIGIN
Query Match 90.2%; Score 1335.4; DB 9; Length 2694;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATCTTTCTTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
Db 284 ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTACCTCTTTTCACTCACT 343
QY 61 GCACCCCTCCATGCTGCTGTACACACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 344 GCACCCCTCCATGCGCTGTATGACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 403
QY 121 ACGGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 404 ATGCAGCGCTCCGGAAATATTGATGCCCATCATATAGGAGTTTCTTAAAGGGAACCTCC 463
QY 181 ACCTTCAGTCCGACACCCATATGCGCGCAACTGCTATAACTCTGCGCACTCTTTGCAATG 240
Db 464 ACCTTCAGTCCGACACCCATATGCGCGCAACTGCTATACTCTGCGCACTCTTTGCAATG 523
QY 241 CATGCAATACTCATTTATTGGACAGGGAAATGATTAATCTTAGTTGCTCGGAGGACTT 300
Db 524 CATGCAATACTCATTTATTGGACAGGGAAATGATTAATCTTAGTTGCTCGGAGGACTT 583
QY 301 GGAGCCACTGCTGTTGAGCTTACTTACCACATACCAGTATGCTCTGATGGGGTGAATT 360
Db 584 GGAGTCACTGCTGTTGAGCTTACTTACCACCAACTGTTATGCTGATGGGGTGAATT 643
QY 361 CAAAGTCAAGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGACAT 420
Db 644 CAAAGTCAAGCAAGAGAAACAAAGTAAAGGAAGTAACTCCCAACTGACCCGGTACAT 703
QY 421 AGCACCCCTAGCCCTACAAAGGAGTACTTCTCAAAACTACATGAAACCTCCGTACC 480
Db 704 GGCACCTCTAGCCCTACAAAGGAGTACTTCTCAAAACTACATGAAACCTCCGTACC 763
QY 481 CATACTCGCTGCTGAGCTTATTTAATACCACCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 764 CATACTCGCTGCTGAGCTTATTTAATACCACCTCACTCGGCTCCATGAGGTCTCAGCC 823
QY 541 CAAAACCTTACTTAAGCTGTTGGATGTCCTCCCTCGCATCTTCAGGCCATACATTTCAATC 600
Db 824 CAAAACCTTACTTAAGCTGTTGGATGTCCTCCCTCGCATCTTCAGGCCATATGTTCAATC 883
QY 601 CCTGTTCTTGAAACAAATGGAAACAACTTTCAGCAGAGAAATAAACACCACTTCGTTTAGTA 660
Db 884 CCTGTTCTTGAAACAAATGGAAACAACTTTCAGCAGAGAAATAAACACCACTTCGTTTAGTA 943
QY 661 GGACCTCTGTTTCCAAATCTGGAAATACCCATACCTCAAACTCACCTGTGTAAATTT 720
Db 944 GGACCTCTGTTTCCAAATCTGGAAATACCCATACCTCAAACTCACCTGTGTAAATTT 1003
QY 721 AGCAATACTATAGACACCAAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACCA 780
Db 1004 AGCAATACTATAGACACCAAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACCA 1063
QY 781 ATAGTCTGCTACCTTCAGGAAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
Db 1064 ATAGTCTGCTACCTTCAGGAAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 1123
QY 841 AATGGCTCTTCAGAACTATGCTTCTCTCATTTAGTCCCTCATTTAGTCCCTCATGACCATCTAC 900

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 2694)	REFERENCE	1 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)	14757826	2 (bases 1 to 2694)
---	-----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------	----------	---------------------

QY 301 GGAGCCACTCTCTGTTGGACTTACTTACCCATACAGATATGTCTGATGGGGGTGAAT 360
DB 584 GGAGTCACTGTCTGTTGGACTTACTTACCCAACTGGTATGTCTGATGGGGGTGAAT 643
QY 361 CAAGGTGAGCAAGAGAGAAAAACAAGTAAAGAAAGCAATCTCCAACTGACCCCGGAGCAT 420
DB 644 CAAGATCAGGCAAGAGAGAAAAACAAGTAAAGAAAGTAACTCTCCAACTCACCCTGGTACAT 703
QY 421 AGACCCCTAGCCCTCAAGAGCAAGTCTCTCAAACTACATGAAACCCCTCCGTACC 480
DB 704 GGCACCTCTAGCCCTCAAGAGCAAGTCTCTCAAACTACATGAAACCCCTCCGTACC 763
QY 481 CATACTCGCTGGTGGAGCTTAAATACCAACCTCTCACTCGGTGTCATGAGGTCTCAGCC 540
DB 764 CATACTCGCTGGTGGAGCTTAAATACCAACCTCTCACTCGGTGTCATGAGGTCTCAGCC 823
QY 541 CAAACCCCTACTAATCTGTTGGATGTGCTCCCTCGCATTCAGGCGCATATTCATC 600
DB 824 CAAACCCCTACTAATCTGTTGGATGTGCTCCCTCGCATTCAGGCGCATATTCATC 883
QY 601 CTGTTCTGAAACATGAGCAACTTTCAGCACAGAAATTAACACCACTTCGTTTGTAGTA 660
DB 884 CTGTTCTGAAACATGAGCAACTTTCAGCACAGAAATTAACACCACTTCGTTTGTAGTA 943
QY 661 GGACCTCTGTTTCCAACTCTGGAATAACCACTCAACCTCACTGTGTAATAATTT 720
DB 944 GGACCTCTGTTTCCAACTCTGGAATAACCACTCAACCTCACTGTGTAATAATTT 1003
QY 721 AGCAATACTATAGACACACAGCTCCCAATGATGAGTGGTGAACACCTCCACACGA 780
DB 1004 AGCAATACTATAGACACACAGCTCCCAATGATGAGTGGTGAACACCTCCACACGA 1063
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTCAGCTCAGCTATCATTTTG 840
DB 1064 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTCAGCTCAGCTATCATTTTG 1123
QY 841 AATGGCTCTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB 1124 AATGGCTCTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
QY 901 ACTGAAACAAGATTTATCAATCATGTGCTAGCTAAGGCCCAACAAACAAGAGTACCCATT 960
DB 1184 ACTGAAACAAGATTTATCAATCATGTGCTAATCTAAGCCCCGCAACAAGAGTACCCATT 1243
QY 961 CTTCTCTTTGTTATCAGACAGAGAGTGTAGGAGATAGTGTCTGCTGCTGCTGCTGCTG 1020
DB 1244 CTTCTCTTTGTTATAGGAGAGAGTGTAGGAGATAGTGTCTGCTGCTGCTGCTGCTG 1303
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
DB 1304 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTTGGAACAG 1363
QY 1081 GTCACTGACTCCCTGCTCAGTCTGCAAGATCAACTTAACTCCCTAGCAGAGTATGCTCT 1140
DB 1364 GTCCGCGACTCCCTGCTCAGTCTGCAAGATCAACTTAACTCCCTAGCAGAGTATGCTCT 1423
QY 1141 CAAATCGAAGAGCTTTAGACTGTCTAACCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1424 CAAATCGAAGAGCTTTAGACTGTCTAACCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1483
QY 1201 GGAGAGAACCTGTTATTTATTTAATCAATCCAGATTTGCTGCTGAGAGAGTTAAAGAA 1260
DB 1484 GGGGAAGAATGCTGTTATTTATTTAATCAATCCAGATTTGCTGCTGAGAGAGTTAAAGAA 1543
QY 1261 ATTCAGATCGAATACATGATAGACAGAGAGCTTCAAAACACCGCAAGCTGGGGCCTC 1320
DB 1544 ATTCAGATCGAATACATGATAGACAGAGAGCTTCAAAACACCGCAAGCTGGGGCCTC 1603
QY 1321 CTCAGCCCAATGAGTCCCTGGTCTCTCCCTTTCTTAGGACCTCTAGCAGCTTAAATTTG 1380
DB 1604 CTCAGCCCAATGAGTCCCTGGTCTCTCCCTTTCTTAGGACCTCTAGCAGCTTAAATTTG 1663
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTTAAACCTCTTGTAGTTTGTCTCTCCAGATTT 1440

DB 1664 CTACTCTCTTTGGACCCCTGTATCTTTAAACCTCTTTGTTAACTTTGTTCTTCCAGATC 1723
QY 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
DB 1724 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCCAA 1764
RESULT 25
AF520521S2 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 6 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.
DEFINITION AF520522
ACCESSION AF520522
VERSION AF520522.1 GI:33411012
KEYWORDS 2 of 2
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 6 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/notes="Caucasian
endogenous_virus: HERV-W"
misc_feature 43..57
/note="splice acceptor site"
CDS 284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17583.1"
/db_xref="GI:33411013"
misc_feature 1705..1720
/note="splice acceptor site"
LTR 1959..2694
ORIGIN
Query Match 90.2%; Score 1335.4; DB 9; Length 2694;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
DB 284 ATGGCCCTCCCTTATCATATTTTCTCTTTACTGTTCTTTTACCCTTTTCACTCTCACT 343
QY 61 GCACCCCTCCATGCTGCTGTGTAACACAGTAGCTCCCTTACAGAGAGTTTCTATGAAGA 120
DB 344 GCACCCCTCCATGCGCTGTATGACACAGTAGCTCCCTTACAGAGAGTTTCTATGGAGA 403
QY 121 AGCGCGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB 404 ATGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 463
QY 181 ACCTTCACCTGCCACACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 240
DB 464 ACCTTCACCTGCCACACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATG 523
QY 241 CATGCAAAATCACTATTATTTGGACAGGGAATAATGATTAATCTTAGTTGCTCTGGAGACTT 300
DB 524 CATGCAAAATCACTATTATTTGGACAGGGAATAATGATTAATCTTAGTTGCTCTGGAGACTT 583
QY 301 GGAGCCACTGCTGTGGACTTACTTACCCTATACAGTATCTCTGATGGGGGTGGAATT 360
DB 584 GGAGTCACTGCTGTGGACTTACTTACCCTATACAGTATCTCTGATGGGGGTGGAATT 643
QY 361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420
DB 644 CAAGATCAGCAAGAGAAAAACAATGTAAAGAAAGTAACTCTCCNACTCACCCGGGTACAT 703
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
DB 704 GGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 763
QY 481 CATACTCGCTGGTGGACCTATTATACACCCCTCACTCGCTCCATGAGTCTCAGCC 540
DB 764 CATACTCGCTGGTGAAGCTATTTAATACACCCCTCACTGGGCTCCATGAGTCTCGGCC 823
QY 541 CAAAACCTACTAACTGTGGATGTGCTCCCTCCCTCACTTCAGGCCATACATTTCAATC 600
DB 824 CAAAACCTACTAACTGTGGATGTGCTCCCTCCCTCACTTCAGGCCATATGTTCAATC 883
QY 601 CTGTTCTTGAAACAATGGAAACAATTTACAGACAGAAATAAACAACCACTTCGTTTGTAGTA 660
DB 884 CTGTTACTTGAAACAATGGAAACAATTTACAGACAGAAATAAACAACCACTTCGTTTGTAGTA 943
QY 661 GGACCTCTGTTTCCCACTGGAAATAACCCATACCTCAAACTCACTGCTGTAAATTT 720
DB 944 GGACCTCTGTTTCCCACTGGAAATAACCCATACCTCAAACTCACTGCTGTAAATTT 1003
QY 721 AGCAATCTATAGACAAACCAAGCTCCCAATGATCAGGTGGGTAAACCTCCCAACGA 780
DB 1004 AGCAATCTATACATAACCAACCACTCCCAATGATCAGGTGGGTAACTCTCCCAACAA 1063
QY 781 ATAGTCTGCCTACCTCAGGAATATTTTTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTG 840
DB 1064 ATAGTCTGCCTACCTCAGGAATATTTTTGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTG 1123
QY 841 AATGGCTCTTCAGAACTATGCTTCTCTCTATCTTAGTGCCCTTATGACCATCTAC 900
DB 1124 AATGGCTCTTCAGAACTATGCTTCTCTCTATCTTAGTGCCCTTATGACCATCTAC 1183
QY 901 ACTGAACAAGATTATACAAATCATGTGCTACCTTAAGCCCCCAACAACAAAGAGTACCCATT 960
DB 1184 ACTGAACAAGATTATACAAATCATGTGCTATCTAGCCCCGCAACAAAGAGTACCCATT 1243
QY 961 CTTCTCTTTTGTATCAGACAGAGTGTAGGACACTAGGTACTGGCAATGGCGATATC 1020
DB 1244 CTTCTCTTTTGTATAGGACAGAGTGTAGGTGCTAGGTACTAGGTACTGGCAATGGCGATATC 1303
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAGAATAAATGGTGCATGGACAG 1080
DB 1304 ACAACCTCTACTCAGTTCTACTACAACTATCTCAGAATAAATGGGACATGGAAACGG 1363
QY 1081 GTCACCTGACTCCCTGGTGCACCTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT 1140

DB 1364 GTCGCCGACTCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1423
QY 1141 CAAATCAAGAGCTTTAGACTTGTACCTACCCCAAGAGGGGAACCTGTTTATTTTFA 1200
DB 1424 CAAATCAAGAGCTTTAGACTTGTACCTACCCCTGAAGAGGGGAACCTGTTTATTTTFA 1483
QY 1201 GGAGAAAGACCTGTTATTATTTAAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA 1260
DB 1484 GGGGAAGAACTGCTGTTATTATTTAAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1543
QY 1261 ATTGAGATCGAAATACAAATGTAGACGAGAGCTTCAAAACACCGAAGCTGGGGCCTC 1320
DB 1544 ATTGAGATCGAAATACAACTAGACGAGAGCTTCAAAACACCTGGACCTGGGGCCTC 1603
QY 1321 CTCAGCCAAATGGATGCTGGTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 1604 CTCAGCCAAATGGATGCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1663
QY 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAAT 1440
DB 1664 CTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTAACTTGTCTCTTCCAGAAATC 1723
QY 1441 GAAGCTGTAAAGCTACAGATGCTCTTCAAAATGGAAACCCCA 1481
DB 1724 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAA 1764

RESULT 26
AF520529S2
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SEGMENT
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AF520529S2
Homo sapiens individual 78 allele A, DNA linear
complete cds, and 3' long terminal repeat, complete sequence.
AF520530
AF520530.1 GI:33411028
2 of 2
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2694)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
2 (bases 1 to 2694)
Mallet,F., Bouton,O. and Oriol,G.
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 78 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="Caucasian
endogenous_virus: HERV-W"
43. .57
/note="splice acceptor site"
284. .1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17587.1"

```

/db_xref="GI:33411029"
/translations="MALPYHIFLFTVLLPSFTTATPPRCMCTSSSPQBFLLWRMQRP
GNTDASYSRLSGKGTFTTATRNCHSATHCMHANTHYWTGMIIPSCPGGLGV
TWCWYPTQTMGGDGVQDAQREKHVEISQLTRVHGTSYSPGVLKSLKHETLR
HTELSLFTNTGLDHEVSAQNTNCWICLPLNFRYPVSIVPVEQNNFSTEINTSV
LVGPLVNLBIHTHSLNTLCPKSNNTVTYNSQCIRWTPPTQICVLPSGIFVCGTSA
LVCLSSGSSSECSFLVLPVNTYIETQDLNVYISKPRNRPVILPFIIVAGVLGALG
TGTGGITSTQFYKYLQSDGDMERVADSLVTLOQLNSLAALVQLNRRALDILLTAE
RGGCLPLGEECYVNSQSGIVTEKYKSIHDRLQRAABELRTGPMGLISQWMPWTLIP
FLGPLAAIILULLFGPICFNLLNVFSSRIEAVLQMEPKMQSKTKIYTRPLDRPASP
RSDVNDIKGTTPPEISAAQPLLRPNSAGSS"
1705..1720
/feature="splice acceptor site"
1959..2694
misc_feature
LTR

```

ORIGIN	Query Match	90.2%	Score 1335.4	DB 9	Length 2694
	Best Local Similarity	93.9%	Pred. No. 0		
	Matches 1390	Conservative 0	Mismatches 91	Indels 0	Gaps 0
Qy	1	ATGGCCCTCCCTTATCATCTTTTCTCTTTACTGTCTCTTACCCCTTTTGGCTCTCACT	60		
Db	284	ATGGCCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCCTTTTCACTCTCACT	343		
Qy	61	GCACCCCTCCATCTGCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTTATGAAGA	120		
Db	344	GCACCCCTCCATCGCGCTGTATGACAGTAGTCTCCCTTACCAAGAGTTTCTTATGGAGA	403		
Qy	121	ACGGGCTTCTCGAAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGAAACTCC	180		
Db	404	ATGCAGGCTCCGGAAATATTGATGCCCCCATCGTATAGGAGTCTTTCTAAGGAAACCCCC	463		
Qy	181	ACCTTCACTGCCACACCCATATGCGCCGCAACTGCTATTAACCTCTGCCACTCTTTGCA	240		
Db	464	ACCTTCACTGCCACACCCATATGCGCCGCAACTGCTATTAACCTCTGCCACTCTTTGCA	523		
Qy	241	CATGCNAATCTCATATTGACAGGGGAAATGATTAACTCTAGTGTCTGCTGGAGGACTT	300		
Db	524	CATGCNAATCTCATATTGACAGGGGAAATGATTAACTCTAGTGTCTGCTGGAGGACTT	583		
Qy	301	GGAGCCACTGTCTGTTGGACTTACTTCAACCCATACCAAGTATGTCTGATGGGGTGGAA	360		
Db	584	GGAGTCACTGTCTGTTGGACTTACTTCAACCCAACTGGTATGTCTGATGGGGTGGAG	643		
Qy	361	CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGG	420		
Db	644	CAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAACTCTCCCAACTCACCCGGGT	703		
Qy	421	AGCACCCCTAGCCCTCAAAAGGACTAGTCTCTCAAAACTACATGAACCCCTCCGTACC	480		
Db	704	GGCACCTTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC	763		
Qy	481	CATACTCGCGTGGTGAGCCTATTAAATACCAACCTCACTCGGCTCCATGAGGTCTCAG	540		
Db	764	CATACTCGCGTGGTAAAGCCTATTAAATACCAACCTCACTCGGCTCCATGAGGTCTCG	823		
Qy	541	CAAAAACCTTACTAACTGTTGGATGTGCGCTCCCGCTGCACTTCAGGCCATACATTTCA	600		
Db	824	CAAAAACCTTACTAACTGTTGGATGTGCGCTCCCGCTGCACTTCAGGCCATATGTTTCA	883		
Qy	601	CCTGTTCTGTGAACAAATGGAACAACTTCAGACAGAGAAATAAACACCACTTCGTTT	660		
Db	884	CCTGTACCTGGAACAAATGGAACAACTTCAGACAGAGAAATAAACACCACTTCGTTT	943		
Qy	661	GGACCTCTGTTTCCAACTCTGGAAATAACCCATACCTCAAAACCTCACTCTGTGTA	720		
Db	944	GGACCTCTGTTTCCAACTCTGGAAATAACCCATACCTCAAAACCTCACTCTGTGTA	1003		
Qy	721	AGCAATACTATAGACACAACACAGCTCCCAATGCAATCAGGTGGGTAAACAACCTCC	780		
Db	1004	AGCAATACTATACACACAACCACTCCCAATGCAATCAGGTGGGTAAACTCTCTCCCA	1063		
Qy	781	ATAGTCTGCCTACCCCTCAGGAATAATTTTGTGTGGTGAACCTCAGCCATCATTTG	840		

RESULT 27	ACCESSION	REFERENCE	JOURNAL
AF520561S2	VERSION	AUTHORS	PUBMED
LOCUS	KEYWORDS		REFERENCE
DEFINITION	SEGMENT		AUTHORS
	SOURCE		TITLE
	ORGANISM		JOURNAL

AF520561S2 2694 bp DNA linear PRI 11-FEB-2004
Homo sapiens individual 71 allele A, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
AF520562
AF520562.1 GI:33411092
2 of 2
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2694)
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G.,
Bonnaud, B., Lucotte, G., Durst, L. and Mandrand, B.
The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
2 (bases 1 to 2694)
Mallet, F., Bouton, O. and Oriol, G.
Direct Submission
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée

FEATURES		d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers		1..2694
source		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/isolate="individual 71 allele A"
		/db_xref="taxon:9606"
		/chromosome="7"
		/map="7q21-q22"
		/sex="male"
		/cell_type="PBMC's"
		/note="African"
		endogenous_virus: HERV-W"
misc_feature		43..57
		/note="splice acceptor site"
CDS		284..1900
		/note="syncytin"
		/codon_start=1
		/product="envelope glycoprotein"
		/protein_id="AAQ17603.1"
		/db_xref="GI:33411093"
		/translation="WALPYHIEFLTVLPSTLTAPPRCMTSSSPYOEFWMQRP
		GNIDAPSVSLSKGPTFTAHMPNRCVHSATLQWANTHYWTKMINPSPGGLGV
		TWCWYFTQGMDSGDQDOAREKHVKEVISQLTRVHGTSSPYKGLDLSLHETLRT
		HRLVSLFNITLGLHEVSAQNPTNCWICLPLNFRPYVISIPVEQNMFEINTISV
		LVGPLVSLNLEIHTHSLNLCVKSNTYITNSOCIRWTPPTQIVCLPSGIFVCGTSA
		YRLNGSSSMCFLSFLVPPMTIYEQDLNLYVSKPRNKRVPILPFVIGVGLGALG
		TGIGTITTSQFYKLSLQBLDMERVADSLVLOQLNSLAUVLQNRRLDLTLAE
		RGCTFLGEECCYVYNQSGIVTEKVEIRDIQRAEELRTPGWLLSQMPPWILP
		FLGPLAAIILLGLGPCIENLLVNFVSSRIEAVLQMEPMQSKTIVRRPLDRPASP
		RSVDNIKGTPPEEISAAQPLLRPNASGS"
misc_feature		1705..1720
		/note="splice acceptor site"
LTR		1959..2694
ORIGIN		
Query Match		90.2%; Score 1335.4; DB 9; Length 2694;
Best Local Similarity		93.9%; Pred. No. 0;
Matches 1390; Conservative		0; Mismatches 91; Indels 0; Gaps 0;
QY	1	ATGGCCCTCCCTTATCATACTTTTCTTTTACTTGTTCTTCTTACCCCTTTTGGCTCTCACT 60
DB	284	ATGGCCCTCCCTTATCATATTTTCTTTTACTTGTTCTTTTACCCCTTTTCACTCTCACT 343
QY	61	GCACCCCTCCATGCTGCTGTACAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB	344	GCACCCCTCCATGCTGCTGTATGACAGTAGCTCCCTTACCAAGAGTTTCTATGGAGA 403
QY	121	ACGGGCTTCTGGAATATTGATGCCCATATATAGAGATTATCTAAGGGAAACTCC 180
DB	404	ATGCAGCGTCCCGGAATATTATGATGCCCATCTGTATAGAGTCTTTCTAAGGGAAACCCC 463
QY	181	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCACTCTTTGGATG 240
DB	464	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCACTCTTTGGCATG 523
QY	241	CATGCAAACTACTATTATGGACAGGAAATAGTTAATCCTAGTTGCTCTGGAGGACTT 300
DB	524	CATGCAAACTACTATTATGGACAGGAAATAGTTAATCCTAGTTGCTCTGGAGGACTT 583
QY	301	GGAGCACTGTCTGTGACCTTACTTCAACCATACCACTATGTCTGATGGGGTGGAAAT 360
DB	584	GGAGTCACTGTCTGTGACCTTACTTCAACCACTGATGTCTGATGGGGTGGAGTT 643
QY	361	CAAGGTTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
DB	644	CAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCAACTCACCCGGGTACAT 703
QY	421	AGCACCTTAGCCCTCAAGAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
DB	704	GGCACTCTAGCCCTCAAGAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 763
QY	481	CATACTCGCCTGGTGAGCCTATTTAATACCAACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Db	764	CATATCGCTGCTGTAAGCCTATTATAACACCCCTCACTGGGTCCATGAGGTCTCGGCC	823
QY	541	CAAAACCTACTAACTGTTGGATGTCCTCCCTGACATTCAGGCCATACATTCAATC	600
Db	824	CAAAACCTACTAACTGTTGGATGTCCTCCCTGACATTCAGGCCATATGTTCAATC	883
QY	601	CCTGTTCTCGAACAATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTA	660
Db	884	CCTGTACTGACCAATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTA	943
QY	661	GGACCTCTTGTTCCTCAATCTGGAATAAACCATACCTCAAACTCCTCACTGTTGTAATTT	720
Db	944	GGACCTCTTGTTCCTCAATCTGGAATAAACCATACCTCAAACTCCTCACTGTTGTAATTT	1003
QY	721	AGCAATACTATAGACACAACCACTCCCAATGCAATCAGGTGGGTAAACCTCCACACAGA	780
Db	1004	AGCAATACTATAGACACAACCACTCCCAATGCAATCAGGTGGGTAACTCTCCACACAA	1063
QY	781	ATAGTCTGCCTACCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG	840
Db	1064	ATAGTCTGCCTACCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG	1123
QY	841	AATGGCTCTTCAGAACTATATGCTGCTTCTCTCATTTCTTAGTGCCCTATGACCATCTAC	900
Db	1124	AATGGCTCTTCAGAACTATATGCTGCTTCTCTCATTTCTTAGTGCCCTATGACCATCTAC	1183
QY	901	ACTGAACAAGATTATTAACAATCATGTGCTACTAAGCCCCCAACAACAAAGAGTAGTACCCATT	960
Db	1184	ACTGAACAAGATTATTAACAATCATGTGCTACTAAGCCCCCAACAACAAAGAGTAGTACCCATT	1243
QY	961	CTTCCTTTTGTATCAGACGAGGAGTGTAGGACAGACTAGGTACTGCGCAITGGCAGTATC	1020
Db	1244	CTTCCTTTTGTATAGGAGCAGGAGTGTAGGTGCTAGGTACTGCGCAITGGCAGTATC	1303
QY	1021	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTGACATGGAACAG	1080
Db	1304	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTGGAACATGGAACGG	1363
QY	1081	GTCAGTACTCCCTGGTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT	1140
Db	1364	GTCGCGGACTCCCTGGTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT	1423
QY	1141	CAAAATCGAAGAGCTTTAGACTGTGTAACCGCAAAAGAGGGGAAACCTGTTATTTTA	1200
Db	1424	CAAAATCGAAGAGCTTTAGACTGTGTAACCGCTGAAGAGGGGAAACCTGTTATTTTA	1483
QY	1201	GGAGAAGAACGCTGTTATTTATGTTAATCAATCCAGATTGTCACTGAGAAAGTTAAAGAA	1260
Db	1484	GGGGAAGAAATGCTGTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA	1543
QY	1261	ATTTCAGAGATCGAATAACAATGTAGACAGAGGAGCTTCAAAAACACCGACGCTGGGCTTC	1320
Db	1544	ATTTCAGATCGAATAACAAGTAGACAGAGGAGCTTCAAAAACACCTGGACCTGGGCTTC	1603
QY	1321	CTCAGCAATAGAACCTCGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
Db	1604	CTCAGCAATAGAACCTCGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1663
QY	1381	TTACTCTCTTTGGACCTGTTATTTAACTCTTGTAAAGTTGTTCTCTTTCCAGAAAT	1440
Db	1664	CTACTCTCTTTGGACCTGTTATTTAACTCTTGTAACTCTTGTAACTTGTCTCTTTCCAGAAAT	1723
QY	1441	GAAGCTGTAAAGCTACAGATGGTCTTTACAAATGGAAACCCCA	1481
Db	1724	GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCACTCCAA	1764

RESULT 28

AX355872

LOCUS

DEFINITION

ACCESSION

AX355872 2930 bp DNA linear PAT 06-FEB-2002

Sequence 1 from Patent WO0204678.

VERSION	AX355872.1	GI:18620523
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 Keith J.C., McCoy J.M. and M.I.S.	
AUTHORS	Methods and compositions for diagnosing and treating preclampsia	
TITLE	and gestational trophoblast disorders	
JOURNAL	Patent: WO 0204678-A 1 17-JAN-2002;	
GENETICS	INSTITUTE, INC. (US)	
FEATURES	Location/Qualifiers	
source	1..2930	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	930..2546	
	/notes="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD23000.1"	
	/db_xref="GI:18620524"	
	/translation="MALPXYHFLFTVLLPSFTLTAAPPKRCMTSSPYQEFELMRMORP TNDVAFYRSLSGSTPTFAHTMPNRYHSALCMHANTHYTGMINPSCPGIGV GIVCWYFTOTGMSDGGVODQAREKHVEVISOLTRVHGTSYPKGLDLSKLHETLR HTPLSLFWTLTGLHEVSAONPTNCWICLPLNFRPVVSI PVEQNNPSTEINTTSV LVGLVSNLEITHSLNLTGVKFSNTYTNISQIRWPTPTOIVCLPSGIFVFCGNSA YRLGSSSESMSELSFLVPPMTIYTEDLLYVVISKPRNKRFPILFVGVAGVLGALG TGGGTTSTQFYKLSQELNGDMERVADELVLTLQDLNSLAALVILONRRALLLTAE RGCTCLFBECECVYVNSQIGIVTEKVEKIDRIQRAEELRNTGPKLLSQWMPWTLF FLGLAAIILLLLFGPCIFNLVNFVSSRIEAVKLQWPKMQSKTKIYRPLDRPASP RSDVNDIKGTTPPEISNAQPLLRPNAGSS"	
ORIGIN	Query Match 90.2%; Score 1335.4; DB 6; Length 2930;	
	Best Local Similarity 93.9%; Pred. No. 0;	
	Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;	
Qy	1	ATGGCCCTCCCTTATCATCTTTCTCTTACTGTTCTTACCTCTTACCCCTTTGGCTCTCCT 60
Db	930	ATGGCCCTCCCTTATCATCTTTCTCTTACTGTTCTTACCTCTTACCCCTTTGGCTCTCCT 989
Qy	61	GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db	990	GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 1049
Qy	121	ACGGGCTTCTCGGAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180
Db	1050	ATGCAGGCTCCCGGAATATTGATGCCCCATCGTATAGGAGTTTCTAAGGGAACTCC 1109
Qy	181	ACCTTCACTGCCGACACCCATATGCCCGGCACTGCTATACCTCTGCCACTTTGGCATG 240
Db	1110	ACCTTCACTGCCGACACCCATATGCCCGGCACTGCTATACCTCTGCCACTTTGGCATG 1169
Qy	241	CATGCAATCTACTATTATGACAGGAAATGATTAATCTAGTCTGCTGGAGGACTT 300
Db	1170	CATGCAATCTACTATTATGACAGGAAATGATTAATCTAGTCTGCTGGAGGACTT 1229
Qy	301	GGAGCCACTGTCTGTGGACTTACTTCAACCATACCATATGCTGATGGGGTGGAAATT 360
Db	1230	GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTT 1289
Qy	361	CAAGGTGAGGCAAGAGAAAAACAAGTAAAGNAGCAATCTCCCAACTGACCCCGGACAT 420
Db	1290	CAAGATCAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 1349
Qy	421	AGCACCCCTAGCCCTCAAAAGGACTAGTCTCTCAAACTACATGAACCTCCGTACC 480
Db	1350	GGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAACTACATGAACCTCCGTACC 1409
Qy	481	CATACCTCGCCCTGGTGGAGCTATTTAATACCAACCTCTACCTCGGCTCCATGAGGTCACGCC 540
Db	1410	CATACCTCGCCCTGGTAAAGCTATTTAATACCAACCTCTACCTCGGCTCCATGAGGTCCTCGCC 1469

Qy	541	CAAAACCCCTACTAATCTGTTGGATGTGCTCCCTCCCTGCACTTCAGGCCATACATTTCAATC 600
Db	1470	CAAAACCCCTACTAATCTGTTGGATATGCTCCCTCCCTGAACTTCAGGCCATATGTTTCAATC 1529
Qy	601	CTGTCTTCTGAACAATGGAACAACCTTCAGCACAGAAATAAACAACCTTCCTGTTTAGTA 660
Db	1530	CTGTACTCTGAACAATGGAACAACCTTCAGCACAGAAATAAACAACCTTCCTGTTTAGTA 1589
Qy	661	GGACCTCTTGTGTTTCAATCTGGAATAAACCCTCAACCTCAACCTCACTGTGTAAATTT 720
Db	1590	GGACCTCTTGTGTTTCAATCTGGAATAAACCCTCAACCTCAACCTCACTGTGTAAATTT 1649
Qy	721	AGCAATCTATATAGACACAAACAGCTCCCAATGCATCAGTGGGTAAACACTCCCAACGGA 780
Db	1650	AGCAATCTATATAGACACAAACAGCTCCCAATGCATCAGTGGGTAAACACTCCCAACGGA 1709
Qy	781	ATAGTCTGCTACCTCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATCTGTTTG 840
Db	1710	ATAGTCTGCTACCTCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATCTGTTTG 1769
Qy	841	AATGGCTCTTCAAGATCTATGTGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
Db	1770	AATGGCTCTTCAAGATCTATGTGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1829
Qy	901	ACTGAAACAAGATTATATACAAATCATGTCTGTAACCTTAAGCCCAACAACAAGAGTACCCATT 960
Db	1830	ACTGAAACAAGATTATATACAAATCATGTCTGTAACCTTAAGCCCAACAACAAGAGTACCCATT 1889
Qy	961	CTTCTCTTTTGTATCAGACGAGGAGTCTAGGACAGTACTAGTACTGGCATGGCAGTATC 1020
Db	1890	CTTCTCTTTTGTATCAGACGAGGAGTCTAGGAGTACTAGTACTGGCATGGCAGTATC 1949
Qy	1021	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAATGTCATGACATGGAACAG 1080
Db	1950	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAATGTCATGACATGGAACAG 2009
Qy	1081	GTCACTGACTCCTGCTGCTACCTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTT 1140
Db	2010	GTCCGCGACTCCTGCTGCTACCTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTT 2059
Qy	1141	CAAAATCGAAGAGCTTTAGACTGCTAACCGCAAAAGAGGGGAACTGTTTATTTTATA 1200
Db	2070	CAAAATCGAAGAGCTTTAGACTGCTAACCGCTGAAAGAGGGGAACTGTTTATTTTATA 2129
Qy	1201	GGAGAAAGAGCTGTTTATTTATGTTTAAATCAATCAGAAATGTCATGAGAAAGTTAAAGAA 1260
Db	2130	GGGAAAGAGCTGTTTATTTATGTTTAAATCAATCCGGAATGTCATGAGAAAGTTAAAGAA 2189
Qy	1261	ATTTCGAGATCGAATACAATGTAGACGAGAGGAGCTTCAAAACACCGAACCTGGGGCTC 1320
Db	2190	ATTTCGAGATCGAATACAACGTAGACGAGAGGAGCTTCAAAACACCGAACCTGGGGCTC 2249
Qy	1321	CTCAGCAATGAGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db	2250	CTCAGCAATGAGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG 2309
Qy	1381	TTACTCTCTTTTGACCCCTGATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAATT 1440
Db	2310	CTACTCTCTTTTGACCCCTGATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAATT 2369
Qy	1441	GAAGCTGTAAGGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db	2370	GAAGCTGTAAGGCTACAAATGGAACCCCAAGATGCAAGTCCAA 2410

RESULT 29
AF208161
LOCUS
DEFINITION Homo sapiens syncytin precursor, mRNA, complete cds.
ACCESSION AF208161
VERSION AF208161.1
GI:6760400
KEYWORDS

SOURCE	Homo sapiens (human)	Db	1050	ATGCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTCTTAAGGAACCC	1109
ORGANISM	Homo sapiens	Qy	181	ACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATAAATCTGCACTCTTTCATG	240
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Db	1110	ACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATAAATCTGCACTCTTTCATG	1169
AUTHORS	1 (bases 1 to 2930)	Qy	241	CATGCAATAACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT	300
	Mi.S., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L., LaValle,E., Tang,X.Y., Edouard,P., Howes,S., Keith,J.C. Jr. and McCoy,J.M.	Db	1170	CATGCAATAACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT	1229
TITLE	Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis	Qy	301	GGAGCCACTGCTGTGGACTTACTTACCACCATACCATGATGCTGATGGGGTGGAAAT	360
JOURNAL	Nature 403 (5771), 785-789 (2000)	Db	1230	GGAGTCACTGCTGTGGACTTACTTACCACCAATCGTATGATGGGGTGGAGTT	1289
MEDLINE	20155476	Qy	361	CAAGTTCAGGCAGAGAAACAAAGTAAGAGAGCAATCTCCCAACTGACCCGGGACAT	420
PUBMED	10693809	Db	1290	CAAGATCAGGCAGAGAAACAAAGTAAGAGAGTAATCTCCCAACTGACCCGGGTACAT	1349
REFERENCE	2 (bases 1 to 2930)	Qy	421	AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAGAACTCCGTACC	480
AUTHORS	Sha.M., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L., LaValle,E., Tang,X., Edouard,P., Howes,S., Keith,J.C. Jr. and McCoy,J.M.	Db	1350	GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAGAACTCCGTACC	1409
TITLE	Direct Submission	Qy	481	CATACCTCGCTGGTGGAGCCCTATTTAATACACCCCTCACTGGCTCCATGAGTCTCAGCC	540
JOURNAL	Submitted (26-NOV-1999) Genetics Institute, 87 Cambridge Park Drive, Cambridge, MA 02140, USA	Db	1410	CATACCTCGCTGGTGGAGCCCTATTTAATACACCCCTCACTGGGCTCCATGAGGTCTCGCC	1469
FEATURES	Location/Qualifiers	Qy	541	CAAAACCCCTAACTGTTGGATGCTCTCCCTGCACCTTCAGGCCCATACATTTCAATC	600
source	1..2930	Db	1470	CAAAACCCCTAACTGTTGGATGCTCTCCCTGCACCTTCAGGCCCATATGTTTCAATC	1529
	/organism="Homo sapiens"	Qy	601	CCTGTTCTCTGAACTGGAACAACCTTCAGCACAGAAATAAACACACCTTCCTGTTTAGTA	660
	/mol_type="mRNA"	Db	1530	CCTGTACCTGAACTGGAACAACCTTCAGCACAGAAATAAACACACCTTCCTGTTTAGTA	1589
	/db_xref="taxon:9606"	Qy	661	GGACCTCTGTTTCCAATCTGGAATAAACCCATACCTCABACCTCACCCTGTGTAATAATTT	720
	/clone="AJ172-28"	Db	1590	GGACCTCTGTTTCCAATCTGGAATAAACCCATACCTCABACCTCACCCTGTGTAATAATTT	1649
	/tissue_type="testis"	Qy	721	AGCAATCTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCACACGA	780
repeat_region	1..2930	Db	1650	AGCAATCTATACATACACAACTCCCAATGCATCAGGTGGGTAAACCTCCACACAA	1709
	/rpt_family="Human endogenous retrovirus HERV-W"	Qy	781	ATAGTCTGCCCTACCTCAGGAATAATTTTGTCTGTGTGCTACTCAGCCCTATCATTTTGTG	840
5'UTR	1..329	Db	1710	ATAGTCTGCCCTACCTCAGGAATAATTTTGTCTGTGTGCTACTCAGCCCTATCATTTTGTG	1769
CDS	930..2546	Qy	841	AATGCTCTTCAGAACTCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC	900
	/note="envelope protein"	Db	1770	AATGCTCTTCAGAACTCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC	1829
	/codon_start=1	Qy	901	ACTGAAACAGATTATTAACAATCATGCTAGTCTAGCCCAACAAACAAAGAGTACCCATT	960
	/product="syncytin precursor"	Db	1830	ACTGAAACAGATTATTAACAATCATGCTAGTCTAGCCCAACAAACAAAGAGTACCCATT	1889
	/protein_id="AAF28334.1"	Qy	961	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGCACTGGGCAATC	1020
	/gb_xref="GI:6760401"	Db	1890	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1949
	/translation="WALPVHIFLFTVLLPSFTLTAPPCCROWTSSSPYOELRWORP GNIDAPYSLSKGPPTFTATHMPNCVHSATLCHNANTHYWTGKMINPSPCGGLGV TWVYFTTGTMSGGVQDQAREKHREKIVISQLTRVHGTSSPYKGLDLSKLHELRIT HTRLVFTNTLTGLHEVAQONRKCICLPLNFRPYSIPVPEQNNPSTELNITSV LVGPLVNLIEIHTSLNLCVKFSNTYTNSQCIKRWTPPTQIVCLPGIFVFCGTSALVGLNGSSMCFPLSPVPMTIYTEQDLYNVIWKPRNKRPVILPFIAGVGLGALG TGGITLTSTQFVKLSQELNGDMERVAADSLVTLQDLNLAVALVQNRALDLTLTAERGTCLFLGEECCYYNQSIGIVTEKVEIKRDIORRAEELRTGPGWLLSOMWPILP FLGPLAAIILLLFGPCIFNLNLFVSSRIEAVKLQMEPKMQSKIKIYRRLDRPASPRSDVNDIKGTPPEITSAAQPLLRPSAGSS"	Qy	901	ACTGAAACAGATTATTAACAATCATGCTAGTCTAGCCCAACAAACAAAGAGTACCCATT	960
	/sig_peptide	Db	1890	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1949
	/mat_peptide	Qy	901	ACTGAAACAGATTATTAACAATCATGCTAGTCTAGCCCAACAAACAAAGAGTACCCATT	960
	/misc_feature	Db	1890	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1949
	/note="furin; cleavage site"	Qy	961	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1020
	/evidence=not_experimental	Db	1890	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1949
	/note="Region: immunosuppressive region"	Qy	961	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1020
	/note="transmembrane-region site"	Db	1890	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1949
	2547..2930	Qy	961	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1020
ORIGIN		Db	1890	CTTCTCTTTTGTATCAGAGGAGTGTAGGACACTAGTCTAGGCACTGGGCAATC	1949
Query Match	90.2%; Score 1335.4; DB 9; Length 2930;	Qy	1021	ACAACTCTCTACCTCAGTTTACTTACAAACTATCTCAAGAAATAAATGGTGACATGGAAACAG	1080
Best Local Similarity	93.9%; Pred. No. 0;	Db	1950	ACAACTCTCTACCTCAGTTTACTTACAAACTATCTCAAGAAATAAATGGGACATGGAAACG	2009
Matches 1390; Conservative	0; Mismatches 91; Indels 0; Gaps 0;	Qy	1081	GTCACTGACTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Qy	1 ATGCCCTCCCTATCATATCTTTCTTTACTTGTCTCTTACCCCTTTTCGCTCTCACT	Db	2010	GTCCGGACTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	2069
Db	930 ATGCCCTCCCTATCATATCTTTCTTTACTTGTCTCTTACCCCTTTTCGCTCTCACT	Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTAAACCGGCTGAAAGAGGGGGAACCTGTTTATTTTAA	1200
Qy	61 GCACCCCTCCATGCTCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA	Db	2070	CAAAATCGAAGAGCTTTAGACTTGTAAACCGGCTGAAAGAGGGGGAACCTGTTTATTTTAA	2129
Db	990 GCACCCCTCCATGCTCTGTATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA	Qy	1201	GGAGAAAGAACGCTGTTTATTTATGTTAATCAATCCAGAAATTTGTCACTGAGAAAGTTAAAGAA	1260
Qy	121 ACGCGGCTCTCTGGAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGAAGTCC	Db	2130	GGGGAAGAAATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA	2189

QY	1261	ATTGAGATCGAATAAATGTAGACGAGAGGCTTCAAAACACGGAACGCTGGGGCTC	1320
Db	2190	ATTGAGATCGAATAAATGTAGACGAGAGGCTTCAAAACACGGAACGCTGGGGCTC	2249
QY	1321	CTCAGCCAATGGATGCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	1380
Db	2250	CTCAGCCAATGGATGCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	2309
QY	1381	TTACTCTCTTTGGACCTGTATCTTTAAACCTCTTTAGTCTTTCCAGAATT	1440
Db	2310	TTACTCTCTTTGGACCTGTATCTTTAAACCTCTTTAGTCTTTCCAGAATT	2369
QY	1441	GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAAACCCCA	1481
Db	2370	GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCCAA	2410
RESULT 30			
ARI177269			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1390; Conservative			
QY	1	ATGGCCCTCCCTTATCATATCTTTCTTTACTGTTCTTTACCCCTTTACCCCTTTTGGCTCTCACT	60
Db	928	ATGGCCCTCCCTTATCATATCTTTCTTTACTGTTCTTTTACCCCTTTTCACTCTCACT	987
QY	61	GCACCCCTCCATGCTGCTGTACACACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA	120
Db	988	GCACCCCTCCATGCTGCTGTATGACCAAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA	1047
QY	121	ACGGGCTTCTGGAATATTGATGCCCCCATATATAGGAGTTTATCTAAGGGAAACTCC	180
Db	1048	ATGCAGCGTCCCGAAATATTGATGCCCCCATGCTATAGGAGTTTCTAAGGGAACTCCC	1107
QY	181	ACCTTCACTGCCCACACCATATGCCCCGCAACTGCTATAACTCTGCCACTTTTGTGATG	240
Db	1108	ACCTTCACTGCCCACACCATATGCCCCGCAACTGCTATCACTCTGCCACTTTTGTGATG	1167
QY	241	CATGCAAACTCATATTATGACAGGGAAATGATTAATCTAGTTGTCTGGAGACTT	300
Db	1168	CATGCAAACTCATATTATGACAGGGAAATGATTAATCTAGTTGTCTGGAGACTT	1227
QY	301	GGAGCCACTGCTGTGACCTTACTTCCACCATACCATAGTATGCTGTAGGGGGTGGAAAT	360
Db	1228	GGAGTCACTGCTGTGTGACCTTACTTCCACCAACTGCTATGCTGTAGGGGGTGGAAAT	1287
QY	361	CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGACAT	420
Db	1288	CAAGTACAGGCAAGAGAAAAACATGTAAGGAAGTAATCTCCCAACTCACCCGGGTACAT	1347
QY	421	AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC	480

LOCUS	BD232464	2946 bp	DNA	linear	PAT 17-JUL-2000			
DEFINITION	Secreted proteins and polynucleotides encoding them.							
ACCESSION	BD232464							
VERSION	BD232464.1	GI:33042234						
KEYWORDS	JP 2002515234-A/2.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
	1 (bases 1 to 2946)							
AUTHORS	Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A.C., Evans,C., Marberg,D., Mi,S. and Treacy,M.							
	Secreted proteins and polynucleotides encoding them							
TITLE	Patent: JP 2002515234-A 2 28-MAY-2002;							
	GENETICS INSTITUTE INC							
COMMENT	OS Homo sapiens (human)							
	PN JP 2002515234-A/2							
	PD 28-MAY-2002							
	PF 17-MAY-1999 JP 2000549638							
	PR 18-MAY-1998 US 09/080478, 20-OCT-1998 US 09/175928 PI							
	KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A COLLINS PI							
	RACIE,							
	PI CHERYL EVANS,DAVID MERBERG,SHA MI,MAURICE TREACY PC							
	C12N15/09,A61K38/00,A61K45/00,A61K48/00,A61P7/00,A61P17/02, PC							
	A61P29/00,							
	PC A61P35/00,A61P37/02,A61P43/00,C07K14/47,C12N5/10,C12P21/02, PC							
	C12Q1/02,							
	PC C12Q1/68,C12N15/00,A61K37/02,C12N5/00							
	CC Secreted proteins and polynucleotides encoding them FH Key							
FEATURES	Location/Qualifiers							
	FT source 1..2946 /organism='Homo sapiens (human)'. /Location/Qualifiers							
ORIGIN	1..2946							
	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"							
Query Match	90.2%; Score 1335.4; DB 6; Length 2946;							
	Best Local Similarity 93.9%; Pred. No. 0; Mismatches 91; Indels 0; Gaps 0;							
Matches 1390;								
Qy	1	ATGGCCCTCCCTTATCATCTTTCTCTTTACTGTTCTCTTACCCCTTTGGCTCTCACT	60					
Db	928	ATGGCCCTCCCTTATCATATATTTCTCTTTACTGTTCTTTTACCCCTCTTCACTCTCACT	987					
Qy	61	GCACCCCTCCATGCTGCTGCTACCAACGATGCTCCCTTACCAGAGTTTCTATGAAGA	120					
Db	988	GCACCCCTCCATGCGCGTGTATGACCAATGCTCCCTTACCAGAGTTTCTATGAAGA	1047					
Qy	121	ACGGGGTCTCTGGAAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAATCC	180					
Db	1048	ATGCAGCGTCCGGAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC	1107					
Qy	181	ACCTTCACTGCCACACCCATATGCCCGCGAATGCTATAACTCTGCCACTTTTGCATG	240					
Db	1108	ACCTTCACTGCCACACCCATATGCCCGCGAATGCTATAACTCTGCCACTTTTGCATG	1167					
Qy	241	CATGCAATATCTATTATTGACAGGGAAATGATTAACTCTAGTTGCTCGGAGGACTT	300					
Db	1168	CATGCAATATCTATTATTGACAGGAAATGATTAACTCTAGTTGCTCGGAGGACTT	1227					
Qy	301	GGAGCCACTGTCTGTGGACTTACTTCAACCATACCAAGTATGTCGTATGGGGTGGAAAT	360					
Db	1228	GGAGTCACTGTCTGTGGACTTACTTCAACCATACCAAGTATGTCGTATGGGGTGGAGTT	1287					
Qy	361	CAAGGTACGGCAAGNAGAAAACAGTAAGCAAGCAATCTCCCAACTGACCCCGGGACAT	420					
Db	1288	CAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACC CGGTACAT	1347					
Qy	421	AGCACCCCTAGCGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC	480					

LOCUS	BD086133	2946 bp	DNA	linear	PAT 27-AUG-2000
DEFINITION	Secreted proteins and polynucleotides encoding them.				
ACCESSION	BD086133				
VERSION	BD086133.1	GI:22631743			
KEYWORDS	JP 2001524490-A/2.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 2946)				
AUTHORS	Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A.C., Evans,C., Merberg,D. and Treacy,M.				
TITLE	Secreted proteins and polynucleotides encoding them				
JOURNAL	Patent: JP 2001524490-A 2 04-DEC-2001;				
COMMENT	GENETICS INSTITUTE INC				
	OS Unidentified				
	PN JP 2001524490-A/2				
	PD 04-DEC-2001				
	PF 17-NOV-1998 JP 200052128				
	PR 21-NOV-1997 US 08/976110, 18-MAY-1998 US 09/080478 PR				
	20-OCT-1998 US 09/175928				
	PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A COLLINS				
	PI RACIE,				
	PI CHERYL EVANS, DAVID MERBERG, MAURICE TREACY				
	PC C07K14/47, C12N15/10, C12N15/09, C12P21/02, C12N5/00, C12N15/00 CC				
	Strandedness: Double;				
	CC Topology: Linear;				
	CC Secreted proteins and polynucleotides encoding them FH Key				
	Location/Qualifiers				
FT	source 1..2946				
FT	/organism='Unidentified'.				
FEATURES	Location/Qualifiers				
source	1..2946				
	/organism='unidentified'				
	/mol_type='genomic DNA'				
	/db_xref='taxon:32644'				
ORIGIN					
Query Match	90.2%;	Score 1335.4;	DB 6;	Length 2946;	
Best Local Similarity	93.9%;	Pred. No. 0;			
Matches 1390;	Conservative 0;	Mismatches 91;	Indels 0;	Gaps 0;	
QY	1	ATGGCCCTCCCTTATCATCTTTCTTTACTGTTCTTTACCCCTTTTGGCTCTCACT	60		
Db	928	ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTACCCCTTTTGGCTCTCACT	987		
QY	61	GCACCCCTCCATGCTGCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA	120		
Db	988	GCACCCCTCCATGCTGCTGTATGACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA	1047		
QY	121	ACGGGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTC	180		
Db	1048	ATGCAGCGTCCCGAAATATTGATGCCCATCTGATAGGAGTCTTTCTAAGGGAACCTC	1107		
QY	181	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATTAACCTGCACTTTTGGCATG	240		
Db	1108	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTTTTGGCATG	1167		
QY	241	CATGCAAAATCATATTATTGGACAGGGAATAATGATTAATCTAGTTGTCTGGAGGACTT	300		
Db	1168	CATGCAAAATCATATTATTGGACAGGGAATAATGATTAATCTAGTTGTCTGGAGGACTT	1227		
QY	301	GGAGCCATGTCTGTTGGCACTTACTTCAACCATACCATAGTATGCTGTATGGGGTGGAA	360		
Db	1228	GGAGTCATGTCTGTTGGCACTTACTTCAACCATACCATAGTATGCTGTATGGGGTGGAA	1287		
QY	361	CAAGGTGAGGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCCACTGACCCCGGGAAT	420		
Db	1298	CAAGATCAGGACAGAGAAAAACATGTAAGAAGTAATCTCCCACTGACCCCGGGAAT	1347		
QY	421	AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAATCATATGAACCTCTGGTACC	480		
Db	1348	GGCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAATCATATGAACCTCTGGTACC	1407		

Db	1405	CTACTCTCTTTGGACCCCTGATCTTTAAACCTCCTTGTGTTAACTTGTGCTCTTCCAGAATC	1464
Qy	1441	GAAGCTGTAAGCTACAGATGGTCTTACAATGGAACCCCA	1481
Db	1465	GAAGCTGTAAGCTACAGATGGTCTTACAATGGAACCCCA	1505
RESULT 34			
AF513360			
LOCUS	1860 bp	mRNA	PRI 05-JUN-2002
DEFINITION	Homo sapiens enverin mRNA, complete cds.		
ACCESSION	AF513360		
VERSION	AF513360.1	GI:21326140	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Alliel, P.M., Perin, J.P., Pierig, R., Nussbaum, J.L., Menard, A. and Rieger, F.		
TITLE	Endogenous retroviruses and multiple sclerosis. Part 2: HERV-7q and its env transcripts		
JOURNAL	C. R. Acad. Sci. III, Sci. Vie 312, 857-863 (1998)		
REFERENCE			
AUTHORS	Alliel, P.M., Perin, J.P., Goudou, D., Bitoun, M., Robert, B. and Rieger, F.		
TITLE	The HERV-W/7q family in the human genome. Potential for protein expression and gene regulation		
JOURNAL	Cell Mol. Biol. 48 (2), 213-217 (2002)		
MEDLINE	21985840		
PUBMED	11990458		
REFERENCE			
AUTHORS	Alliel, P.M., Goudou, D., Perin, J.P. and Rieger, F.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-MAY-2002) U-488, INSERM, 80, rue du General Leclerc, Le Kremlin-Bicetre 94270, France		
FEATURES			
source	Location/Qualifiers		
	1..1860		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="7"		
	/map="7q21"		
	/tissue_type="placenta"		
	59..1675		
	/codon_start=1		
	/product="enverin"		
	/protein_id="AAM47599.1"		
	/db_xref="GI:21326141"		
	/translation="MALPYHI FLPTVLLPSFTLTAPPPRCMTSSSPYQFLNRMQRP GNDAPSYRLSKGTPTTFAHTMPCNYHSALCMHANTHYWTGRMINSPCGGLGV TVCYTYFTQMSDGGVQDQAREKHVKEVISQLTRVHGTSPPYKGLDLKSLHETLRT HTRLSVLTNTTGLHVSQNPNCWICLPNRPYVSI PVPEQMNFNTEINTTSV LVGPLVNLHTSNTLCVKFSNTVTNCSOCIRWVTPPTQIVCLPSGIPFVCGTSA YRLGNSSEKCFLSFLVPPMTIYEQDLNLYVTSKPNKRVPLLPVIGAGVLGALG YRLGNTTSCFYKLSQELNGDMERVADESLTQDLNLSAAVLQNRALLDLTNE RGCTFLGEECCYYNQSGIVTEKVEIRDRIOORRAEELRNGPWLLSQMPFWPLP FLGPLAAIILLLFPGPCIFNLLNVFVSRIEAVKLQMEPKMQSKTKIYRRLDRPASP RSDVNDIKGTPPEISAAQPLLRNPSAGSS"		
CDS			
	/codon_start=1		
	/product="enverin"		
	/protein_id="AAM47599.1"		
	/db_xref="GI:21326141"		
	/translation="MALPYHI FLPTVLLPSFTLTAPPPRCMTSSSPYQFLNRMQRP GNDAPSYRLSKGTPTTFAHTMPCNYHSALCMHANTHYWTGRMINSPCGGLGV TVCYTYFTQMSDGGVQDQAREKHVKEVISQLTRVHGTSPPYKGLDLKSLHETLRT HTRLSVLTNTTGLHVSQNPNCWICLPNRPYVSI PVPEQMNFNTEINTTSV LVGPLVNLHTSNTLCVKFSNTVTNCSOCIRWVTPPTQIVCLPSGIPFVCGTSA YRLGNSSEKCFLSFLVPPMTIYEQDLNLYVTSKPNKRVPLLPVIGAGVLGALG YRLGNTTSCFYKLSQELNGDMERVADESLTQDLNLSAAVLQNRALLDLTNE RGCTFLGEECCYYNQSGIVTEKVEIRDRIOORRAEELRNGPWLLSQMPFWPLP FLGPLAAIILLLFPGPCIFNLLNVFVSRIEAVKLQMEPKMQSKTKIYRRLDRPASP RSDVNDIKGTPPEISAAQPLLRNPSAGSS"		
ORIGIN			
Query Match	90.1%	Score 1333.8;	DB 9; Length 1860;
Best Local Similarity	93.8%	Pred. No. 0;	
Matches 1389;	Conservative	0; Mismatches	92; Indels 0; Gaps 0;
Qy	1	ATGCGCCCTCCCTTATCATCTTCTCTTACTGTCCTTACCCCTTCCGCTCTCACT	60
Db	59	ATGCGCCCTCCCTTATCATCTTCTCTTACTGTCCTTACCCCTTCTTCACTCTCACT	118
Qy	61	GCACCCCTCCATGCTGCTATCAACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA	120
Db	119	GCACCCCTCCATGCTGCTATCAACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA	178
Qy	121	ACGCGCTTCTTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC	180
Db	179	ATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGAAACCC	238
Qy	181	ACCTTCACTGCCACACCCATATGCCCGCACTGCTATAACTCTGCCACTCTTTGCATG	240
Db	239	ACCTTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCCACTCTTTGCATG	298
Qy	241	CATGCAAAATACCTATTATTGGACAGGAAATGATTAACTCTAGTTGTCTCGGAGACTT	300
Db	299	CATGCAAAATACCTATTATTGGACAGGAAATGATTAACTCTAGTTGTCTCGGAGACTT	358
Qy	301	GGAGCCACTGTCTGTGGACTTTACTTCAACCATACCAGTATGTCTGATGGGGTGGAAAT	360
Db	359	GGAGTCACTGTCTGTGGACTTTACTTCAACCATACCAGTATGTCTGATGGGGTGGAGTT	418
Qy	361	CAAGGTGAGGCAAGAGAAAACAAAGTAAGAGAGCAATCTCCCACTGAGCCCGGGACAT	420
Db	419	CAAGATCAGGCAAGAGAAAACAAAGTAAGAGAGTAACTCTCCCACTGAGCCCGGGTACAT	478
Qy	421	AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCTCCGATACC	480
Db	479	GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGATACC	538
Qy	481	CATACCTCGCTGTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC	540
Db	539	CATACCTCGCTGTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCGGCC	598
Qy	541	CAAAAACCTTAACTGTGTGGATGTCCTCCCTCGCATCTTCAAGGCATACATTTCAATC	600
Db	599	CAAAAACCTTAACTGTGTGGATGTCCTCCCTCGCATCTTCAAGGCATATGTTTCAATC	658
Qy	601	CTGTGTTCTGAAACAATGGAACAACTTTCAGCAGAGAAATAAACACCCTCCGTTTTAGTA	660
Db	659	CTGTGTTCTGAAACAATGGAACAACTTTCAGCAGAGAAATAAACACCCTCCGTTTTAGTA	718
Qy	661	GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAAACCTCACTGTGTAATAATTT	720
Db	719	GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAAACCTCACTGTGTAATAATTT	778
Qy	721	AGCAATATATACACCAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA	780
Db	779	AGCAATATATACACCAACAGCTCCCAATGATCAGGTGGGTAACTCTCCACACCAA	838
Qy	781	ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG	840
Db	839	ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG	898
Qy	841	AATGGCTCTTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCTTATGACCATCTAC	900
Db	899	AATGGCTCTTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCTTATGACCATCTAC	958
Qy	901	ACTGAAACAAGATTTATACAAATCATGTCGTACCTAAGCCCCCAACAAAAGAGTACCCATT	960
Db	959	ACTGAAACAAGATTTATACAAATTTATGTCATATCTAAGCCCCCAACAAAAGAGTACCCATT	1018
Qy	961	CTTCTCTTTGTTTATCAGACGAGGTGCTAGGCAGAGTACTGTCATTGGCAGTATC	1020
Db	1019	CTTCTCTTTGTTTATGAGGACGAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGTGCTATC	1078
Qy	1021	ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAACAG	1080
Db	1079	ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAACAG	1138
Qy	1081	GTCACTGACTCCCTGGTCACTTGCAGAGATCAACTTAACTCCCTTAGCAGAGTAGTCTCT	1140
Db	1139	GTGCGGACTCCCTGGTCACTTGCAGAGATCAACTTAACTCCCTTAGCAGAGTAGTCTCT	1198
Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAGAGGGGAACTGTTTATTTTAA	1200
Db	1199	CAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAAAGAGGGGAACTGTTTATTTTAA	1258

QY	1201	GGAGAAGAACGCTGTTATTTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA	1260	Db	511	ATGAGCGTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC	570
Db	1259	GGGGAAGAATGCTGTTACTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTAAAGAA	1318	QY	181	ACCTTCACTGCCACACCCATATGCCCCGCAATGCTTAACCTCTGCCACTCTTTGGATG	240
QY	1261	ATTCCAGATCGAATCAATGTAGACGAGGAGCTTCAAAAACACGAACGCTGGGGCCTC	1320	Db	571	ACCTTCACTGCCACACCCATATGCCCCGCAATGCTTAACCTCTGCCACTCTTTGGATG	630
Db	1319	ATTCCAGATCGAATCAACGTAGACGAGGAGCTTCAAAAACACCTGGACCTGGGGCCTC	1378	QY	241	CATGCAATCTCAATATTGACAGAGGAAATGATTAATCCTAGTTGCTCTGGAGGACTT	300
QY	1321	CTCAGCAATGATGCGCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATG	1380	Db	631	CATGCAATCTCAATATTGACAGGAAATGATTAATCCTAGTTGCTCTGGAGGACTT	690
Db	1379	CTCAGCAATGATGCGCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATG	1438	QY	301	GGAGCCACTGTCTTTGGACTTACTTCAACCATACAGTATGTCTGATGGGGTGGAAAT	360
QY	1381	TTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTAGTTTGTCTCTCCAGAAAT	1440	Db	691	GGAGTCACTGTCTTTGGACTTACTTCAACCATACAGTATGTCTGATGGGGTGGAGTT	750
Db	1439	CTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTAGTTTGTCTCTCCAGAAAT	1498	QY	361	CAAGGTGAGCAAGAGAAACAAAGTAAGGAAGCAATCTCCCAACTGACCCGGGACAT	420
QY	1441	GAAAGTGTAAAGTACAGATGCTTTACAAATGGAAACCCCA	1481	Db	751	CAAGATCAGGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGACAT	810
Db	1499	GAAAGTGTAAAGTACAAATGGAGCCCAAGATGCAATCCAA	1539	QY	421	AGCACCTCTAGCCCTACAAAGAGTACTTCTCTCAAAAATACATGAAACCTCCGTACC	480
RESULT 35				Db	811	GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACC	870
BD221827		2055 bp DNA linear PAT 17-JUL-2003		QY	481	CATCTCGCTGTGGAGCTTATTAATACCACTCTCTCGGCTCCATGAGTCTCAGCC	540
LOCUS		Nucleic sequence and deduced protein sequence family with human		Db	871	CATCTCGCTGTGGAGCTTATTAATACCACTCTCTCGGCTCCATGAGTCTCAGCC	930
DEFINITION		endogenous retroviral motifs, and their uses.		QY	541	CAAAACCTCTACTACTGTTGGATGCTCCCTCCCTGCACTTTCAGGCCATACATTTCAATC	600
ACCESSION		BD221827.1 GI:33031597		Db	931	CAAAACCTCTACTACTGTTGGATGCTCCCTCCCTGCACTTTCAGGCCATACATTTCAATC	990
VERSION		JP 2002518051-A/22.		QY	601	CCTGTTCTGCAACCAATGGAACTTTCAGCACAGAAATAAAACACACTTCCTGTTTACTA	660
KEYWORDS		Homo sapiens (human)		Db	991	CCTGTTCTGCAACCAATGGAACTTTCAGCACAGAAATAAAACACACTTCCTGTTTACTA	1050
SOURCE		Homo sapiens		QY	661	GGACCTCTGTTTCCAAATCTCGAAATAACCATACCTCAAACTCAGCTGTGTAAATTT	720
ORGANISM		Homo sapiens		Db	1051	GGACCTCTGTTTCCAAATCTCGAAATAACCATACCTCAAACTCAGCTGTGTAAATTT	1110
REFERENCE		Alliel,P.M., Perin,J.P. and Rieger,F.		QY	721	AGCAATCTATAGACACACACAGCTCCCAATGCACTCAGTGGTGAACACTCCCAACAGA	780
AUTHORS		1 (bases 1 to 2055).		Db	1111	AGCAATCTATACATACACACCACTCCCAATGCACTCAGTGGTGAACACTCCCAACAGA	1170
TITLE		Nucleic sequence and deduced protein sequence family with human		QY	781	ATAGTCTGCTACCTCAGGAAATATTTTGTCTGTGTACTCTCAGCTTATCATTTGTTG	840
JOURNAL		Patent: JP 2002518051-A 22 JUN-2002;		Db	1171	ATAGTCTGCTACCTCAGGAAATATTTTGTCTGTGTACTCTCAGCTTATCATTTGTTG	1230
COMMENT		INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE		QY	841	AATGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTTATGACCATCTAC	900
		OS Homo sapiens (human)		Db	1231	AATGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCCCTTATGACCATCTAC	1290
		PN JP 2002518051-A/22		QY	901	ACTGAAACAGATTTATACAGTATGTCATATCTAAGCCCAACAAAGAGTACCCATT	960
		PD 25-JUN-2002		Db	1291	ACTGAAACAGATTTATACAGTATGTCATATCTAAGCCCAACAAAGAGTACCCATT	1350
		PF 23-JUN-1999 JP 2000556036		QY	961	CTTCTCTTTGTTATCAGAGCAGGAGTGTCTAGGACAGACTAGGTAATGGCAGTATC	1020
		PR 23-JUN-1998 FR 98/07920		Db	1351	CTTCTCTTTGTTATAGGAGCAGGAGTGTCTAGGACAGACTAGGTAATGGCAGTATC	1410
		PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC		QY	1021	ACAACTCTCTACAGTCTTCTACAACTATCTCAAGAAATAAATGTCACATGGAACAG	1080
		C12N15/09, A01K67/027, A61K31/711, A61K39/21, A61K48/00, A61P21/00, PC		Db	1411	ACAACTCTCTACAGTCTTCTACAACTATCTCAAGAAATAAATGTCACATGGAACAG	1470
		A61P25/00,		QY	1081	GTCACTGTCTCTGCTGCTCAGCTTCAAGATCAACTTAATCTCCTAGCAGCAGTAGTCTTT	1140
		PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC Nucleic		Db	1471	GTCCGCGACTCTCTGCTCAGCTTCAAGATCAACTTAATCTCCTAGCAGCAGTAGTCTTT	1530
		sequence and deduced protein sequence family with CC		QY	1141	CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTTATTTTA	1200
		human endogenous		Db	1531	CAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAAAGAGGGGAACTGTTTATTTTA	1590
		CC retroviral motifs, and their uses		QY	1201	GGAGAAGAACGCTGTTATTTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA	1260
		PH Key Location/Qualifiers		Db			
		FT CDS (1) . . (2055).		QY			
FEATURES		Location/Qualifiers					
source		1..2055					
		/organism="Homo sapiens"					
		/mol_type="genomic DNA"					
		/db_xref="taxon:9606"					
ORIGIN							
Query Match		90.1%; Score 1333.8; DB 6; Length 2055;					
Best Local Similarity		93.8%; Pred. No. 0;					
Matches 1389; Conservative		0; Mismatches 92; Indels 0; Gaps 0;					
QY	1	ATGCGCCCTCCCTTATCATCTTTTCTCTTACTGTTCTTCTTACCCCTTTGCTCTCACT	60				
Db	391	ATGCGCCCTCCCTTATCATCTTTTCTCTTACTGTTCTTCTTACCCCTTTTCACTCTCACT	450				
QY	61	GCACCCCTCCAGTCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAGA	120				
Db	451	GCACCCCTCCAGTCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAGA	510				
QY	121	ACGGCGCTTCTCGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC	180				


```
QY 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTAAAGTTTGTCTCTCCAGAAATT 1440
Db 1771 CTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTAAAGTTTGTCTCTCCAGAAATC 1830

QY 1441 GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA 1481
Db 1831 GAAGCTGTAAACTACAAATGAGCCCAAGATGCAATCCAA 1871

RESULT 37
BD221806
LOCUS BD221806 2599 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses.
ACCESSION BD221806.1 GI:33031576
VERSION JF 2002518051-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2599)
AUTHORS Alliel,P.M., Perin,J.P. and Rieger,F.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: JP 2002518051-A 1 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Homo sapiens (human)
PN JP 2002518051-A/1
PD 25-JUN-2002
PF 23-JUN-1999 JP 2000556036
PI 23-JUN-1998 FR 98/07920
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
A61P25/00,
PC A61P37/06,C07K14/15,C12Q1/70,C12Q1/68,C12N15/00 CC Nucleic
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1..2599
FT /organism='Homo sapiens (human)'.
FEATURES
source
Location/Qualifiers
1..2599
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 90.1%; Score 1333.8; DB 6; Length 2599;
Best Local Similarity 93.8%; Pred. NO. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATATCTTTCTCTTACTGTGTTCTTACCCCTTTTCGCTCTCACT 60
Db 929 ATGGCCCTCCCTTATCATATTTTCTTTACTGTGTTCTTTACCCCTTTTCACTCTCACT 988
QY 61 GCACCCCTCCATGCTGTGTACAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
Db 989 GCACCCCTCCATGCGCGCTGTATGACAGTAGCTCCCTTTACCAAGAGTTTCTATGGAGA 1048
QY 121 ACGCGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
Db 1049 ATGACGGTCCCGGAATATTGATGCCCATCATGTATAGAGTCTTTCTAAGGAAACCCCC 1108
QY 181 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
Db 1109 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 1168
QY 241 CATGCAAACTACTATTATGGACAGGGAAATGATTAATCCTAGTTGTCTCTGGAGACTT
Db 1169 CATGCAAACTACTATTATGGACAGGGAAATGATTAATCCTAGTTGTCTCTGGAGACTT 1228
```

```
QY 301 GGACCACTGTCTGTGGACTTACTTCCACCATAACCACTATGTCTGATGGGGTGGAAATT 360
Db 1229 GGAGTCACTGTCTGTGGACTTACTTCCACCATAACCACTATGTCTGATGGGGTGGAGTT 1288
QY 361 CAAGGT CAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGGACAT 420
Db 1289 CAAGAT CAGGCAAGAGAAAAAACAATGTAAAAAGAGTAATCTCCCAACTCACCCGGGTACAT 1348
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCCGTTACC 480
Db 1349 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCGTTACC 1408
QY 481 CATACTCGCTGTGTGAGCCTATTTAATACCAACCCCTCACTCGGCTCCATAGAGTCTCAGCC 540
Db 1409 CATACTCGCTGTGTGAGCCTATTTAATACCAACCCCTCACTCGGCTCCATAGAGTCTCGGCC 1468
QY 541 CAAAACCCCTACTAATCTGTGTGATGTGCTCCCTCCCTGACCTTCAGGCCCATATGTTTCAATC 600
Db 1469 CAAAACCCCTACTAATCTGTGTGATATGCTCCCTCCCTGAACTTCAGGCCCATATGTTTCAATC 1528
QY 601 CCTGTTCTGAACAATGGAACAACTTCAGCACAGAAATAAACAACACTTCCGTTTGTAGTA 660
Db 1529 CCTGTACTGTAACAATGGAACAACTTCAGCACAGAAATAAACAACACTTCCGTTTGTAGTA 1588
QY 661 GGACCTCTGTTTCCAATCTGGAATAAACCCATACCTCAAACTCACCTGCTGTGTAATAATT 720
Db 1589 GGACCTCTGTTTCCAATCTGGAATAAACCCATACCTCAAACTCACCTGCTGTGTAATAATT 1648
QY 721 AGCAATACTATAGACAAACCCAGCTCCCAATGCAATGCAATGAGTGGGTAACTCCCTCCCAACA 780
Db 1649 AGCAATACTATAGACAAACCCAGCTCCCAATGCAATGCAATGAGTGGGTAACTCCCTCCCAACA 1708
QY 781 ATAGTCTGCTTACCTCAGGAATAATTTTCTGTGTGTACTCTCAGCCCTATCATTTGTTG 840
Db 1709 ATAGTCTGCTTACCTCAGGAATAATTTTCTGTGTGTACTCTCAGCCCTATCGTTGTTG 1768
QY 841 AATGCTCTTCAGAAATCTATGTGTTCTCTCAATTTAGTGTGCCCCCTATGACCACTAC 900
Db 1769 AATGCTCTTCAGAAATCTATGTGTTCTCTCAATTTAGTGTGCCCCCTATGACCACTAC 1828
QY 901 ACTGAACAAGATTTTATACAATCATGTGCTAGCTTAAGCCCAACAAAGAGTACCCATT 960
Db 1829 ACTGAACAAGATTTTATACAGTTATGTCTATCTAAGCCCAACAAAGAGTACCCATT 1888
QY 961 CTTCCCTTTTGTATCAGAGCAGGAGTGTCTAGGCAGACTAGGTACTGCGAATGGCAGTATC 1020
Db 1889 CTTCCCTTTTGTATAGGAGCAGGAGTGTCTAGGTGCACTAGGTACTGCGAATGGCAGTATC 1948
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
Db 1949 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTTGGAACAGG 2008
QY 1081 GTCAGTACTGCTGCTGTCACCTTGCAGATCAACTTAACTTAACTCCCTAGCAGCAGTACTCCT 1140
Db 2009 GTGCGCAGTCTCCCTGCTCACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTACTCCT 2068
QY 1141 CAAAATCAAGAGCTTTTACACTTGTAAACCGCAAAAGAGGGGGAACCTGTTTATTTTA 1200
Db 2069 CAAAATCAAGAGCTTTTACACTTGTAAACCGCTGAAGAGGGGGAACCTGTTTATTTTA 2128
QY 1201 GGAGAAAGACCGTGTATTATTATCAATCCAGAAATGTGTCATGAGAAAGTTAAAGNA 1260
Db 2129 GGGAAGAAATGCTGTTATTATTATCAATCCGGAATCGTCATCGAGAAAGTTAAAGNA 2188
QY 1261 ATTCAGATCGAATACATATGATGACAGAGGAGCTTCAAAACACCGAAGCTGGGCGCTC 1320
Db 2189 ATTCAGATCGAATACATGATGACAGAGGAGCTTCAAAACACCTGGACCCCTGGGCGCTC 2248
QY 1321 CTCAGCAATGATGCCCTGGGTTCTCCCTTTCTTTAGGACCTCTTAGCAGCTCTAATATTG 1380
Db 2249 CTCAGCAATGATGCCCTGGGTTCTCCCTTTCTTTAGGACCTCTTAGCAGCTCTAATATTG 2308
QY 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTTGTAGTTGTCTCTCCAGAAATT 1440
```


Db 2309 CTACTCTCTTGGACCTGATCTTAACTCTCTTGTAACTTGTCTCTCCAGATC 2368
 QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAATGGAACCCCA 1481
 Db 2369 GAAGCTGTAAGCTACAGATGGTCTTACAATGGAACCCCA 2409

RESULT 38
 AX007978 2599 bp DNA linear PAT 06-SEP-2000
 LOCUS Sequence 1 from Patent WO9967395.
 DEFINITION AX007978
 ACCESSION AX007978
 VERSION AX007978.1 GI:9995675
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Perin,J.P., Rieger,F. and Alliel,P.M.
 Nucleic sequence and deduced protein sequence family with human
 endogenous retroviral motifs, and their uses
 Patent: WO 9967395-A 1 29-DEC-1999;
 INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
 FRANCOIS (FR); ALLIEL PATRICK M (FR)

FEATURES
 Location/Qualifiers
 1. 2599
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 90.1%; Score 1333.8; DB 6; Length 2599;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATCTTTCTCTTACTGTTCTTACCCCTTTGGCTTCACCT 60
 Db 929 ATGGCCCTCCCTTATCATCTTTCTCTTACTGTTCTTACCCCTTTTTCACCTCTC 98
 QY 61 GCACCCCTCCATGCTGTACAAACAGTAGCTCCCTTACCAAGAGTTCTTATGAAGA 120
 Db 989 GCACCCCTCCATGCTGTATGACAGTAGCTCCCTTACCAAGAGTTCTTATGAAGA 1048
 QY 121 ACGGGCTTCTGTGAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAAATCC 180
 Db 1049 ATGACGGTCCCGAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAAATCC 1108
 QY 181 ACCTTCACTGCCACACCATATGCGCGGCACTGTCTAATCTGCGACTCTTTGCATG 240
 Db 1109 ACCTTCACTGCCACACCATATGCGCGGCACTGTCTAATCTGCGACTCTTTGCATG 1168
 QY 241 CATCAATATCTAATTTGACAGGAAATGATTAATCTAGTCTCTGGAGGACTT 300
 Db 1169 CATCAATATCTAATTTGACAGGAAATGATTAATCTAGTCTCTGGAGGACTT 1228
 QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCATATGCTGTATGGGGTGGAAAT 360
 Db 1229 GGAGTCACTGTCTGTGGACTTACTTCAACCATACCATATGCTGTATGGGGTGGAGTT 1288
 QY 361 CAAGGTGAGGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCAACTGACCCGGGGACAT 420
 Db 1289 CAAGTCAAGGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCAACTGACCCGGGTACAT 1348
 QY 421 AGCACCTTAGCCCTACAGAGGACTAGTCTCTCAAACTACATGAACCTCCGTACC 480
 Db 1349 GGCACCTTAGCCCTACAGAGGACTAGTCTCTCAAACTACATGAACCTCCGTACC 1408
 QY 481 CATACTCGCTGGTGAGCTATTTAATACCACTCTCACTCGGCTCCATGAGGTCTCAGCC 540
 Db 1409 CATACTCGCTGGTGAGCTATTTAATACCACTCTCACTCGGCTCCATGAGGTCTCGCC 1468

QY 541 CAAACCCCTCTAATCTGTTGGATGCTCCCTGCACTTCCAGGCATACATTTCAATC 600
 Db 1469 CAAACCCCTCTAATCTGTTGGATGCTCCCTGCACTTCCAGGCATACATTTCAATC 1528
 QY 601 CTTGTTCTGAACTGGAACAACTTCAAGCAGAGAAATAACACACACTTCCGTTTGTAGTA 660
 Db 1529 CTTGTTCTGAACTGGAACAACTTCAAGCAGAGAAATAACACACACTTCCGTTTGTAGTA 1588
 QY 661 GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCTCACTGTGTAAATTT 720
 Db 1589 GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCTCACTGTGTAAATTT 1648
 QY 721 AGCAATATATAGACAAACACAGCTCCCAATGATCAGGTGGTAAACACTCCCAACAGTA 780
 Db 1649 AGCAATATATAGACAAACACAGCTCCCAATGATCAGGTGGTAAACACTCCCAACAGTA 1708
 QY 781 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 840
 Db 1709 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 1768
 QY 841 AATGGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 Db 1769 AATGGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1828
 QY 901 ACTGAACAAGATTTATACAAATCATGTCCTAAGCCCAACAAACAAAGAGTACCCATT 960
 Db 1829 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCAACAAACAAAGAGTACCCATT 1888
 QY 961 CTTCTCTTTTGTATCAGACAGAGTGTCTAGGACACTAGGTACTGGCATTGGCAGTATC 1020
 Db 1889 CTTCTCTTTTGTATCAGACAGAGTGTCTAGGACACTAGGTACTGGCATTGGCAGTATC 1948
 QY 1021 ACACCT 1080
 Db 1949 ACACCT 2008
 QY 1081 GTCACTGACTCTCCCTGCTCACTTTCAGAGATCAACTTAACTCCCTAGCAGCAGTAGTCT 1140
 Db 2009 GTCCGCGACTCTCCCTGCTCACTTTCAGAGATCAACTTAACTCCCTAGCAGCAGTAGTCT 2068
 QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAGCCCAAGAGGGGGGAACTGTTTATTTTAA 1200
 Db 2069 CAAATCGAAGAGCTTTAGACTTGTCTAAGCCCAAGAGGGGGGAACTGTTTATTTTAA 2128
 QY 1201 GGAAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATGTTCTACTGAGAAAGTTAAAGAA 1260
 Db 2129 GGAAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATGTTCTACTGAGAAAGTTAAAGAA 2188
 QY 1261 ATTCGAGATCGAATACAATGTAGACAGAGGAGCTTCAAAACACCGAACCTGGGGCTC 1320
 Db 2189 ATTCGAGATCGAATACAATGTAGACAGAGGAGCTTCAAAACACCGAACCTGGGGCTC 2248
 QY 1321 CTCAGCAATGAGTGGCTGGGTTCTCCCTCTTATAGGACCTCTAGCAGCTCTAATATTG 1380
 Db 2249 CTCAGCAATGAGTGGGTTCTCCCTCTTATAGGACCTCTAGCAGCTCTAATATTG 2308
 QY 1381 TTAATCTCTTCTGGACCTCTGTTTAACTCTCTTAACTCTCTTAACTCTCTTAACTCTCT 1440
 Db 2309 TTAATCTCTTCTGGACCTCTGTTTAACTCTCTTAACTCTCTTAACTCTCTTAACTCTCT 2368
 QY 1441 GAAGCTGTAAGCTACAGATGGTCTTCAAACTGGAACCCCA 1481
 Db 2369 GAAGCTGTAAGCTACAGATGGTCTTCAAACTGGAACCCCA 2409

RESULT 39
 AF520481S2
 LOCUS Homo sapiens individual 147 allele A, envelope glycoprotein gene,
 DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
 AF520482
 ACCESSION AF520482
 VERSION AF520482.1
 KEYWORDS GI:33410932

Db 1604 CTCAGCAGTATGATGCGCTGGATTCTCCCTCTTAGGACCTCTAGCAGCTATAATATTG 1663

Qy 1381 TTACTCTCTTTGGACCTGATCTTTAACTCTCTTAAAGTTTGTCTTCCAGAAATT 1440

Db 1664 CTACTCTCTTTGGACCTGATCTTTAACTCTCTTAAAGTTTGTCTTCCAGAAATC 1723

Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481

Db 1724 GAAGCTGTAAACTACAAATGGAGCCCAAGATGAGTCCAA 1764

RESULT 40

AF520483S2 2694 bp DNA linear PRI 11-FEB-2004

LOCUS Homo sapiens individual 147 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.

AF520484

AF520484.1 GI:33410936

2 of 2

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2694)

Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.

The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology

Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

14757826

2 (bases 1 to 2694)

Mallet, F., Bouton, O. and Oriol, G.

Direct Submission

Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France

Location/Qualifiers

1. .2694

/organism="Homo sapiens"

/mol_type="genomic DNA"

/isolate="individual 147 allele B"

/db_xref="taxon:9606"

/chromosome="7"

/map="7q21-q22"

/sex="female"

/cell_type="PBMC's"

/note="endogenous_virus: HERV-W"

43. 57

/note="splice acceptor site"

misc_feature

284. .1900

/note="syncytin"

/codon_start=1

/product="envelope glycoprotein"

/protein_id="AA017564.1"

/db_xref="GI:33410937"

/translation="MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQBFELRMQRP GNIDAPSYRSLGKSTPTTAHTHPRNCYHSATLCMHANTHYWTGRMINSFPGGLV TVCWYTFQTSMDGGVQDAEKHKVEVISQLTRVHGTSSPYKGLDLKSLHETLR HTRVLSPNLTTLGLHVSQNPNTNCWICLPNRPVYSIPVPEOWNPSEINTTSSV LVGLVSNLEITHPSLTCKVFNSTVTTNSQICRWVTPPTQIVCLPSGIFVFCGTA YRLVNSGSMCFUSFLVPWMTIYEQDLISYVSKPRKRPVLPFVIGAVILGALG TGIGGITTSQFYKLSQELNGDMERVADSLVTLDQNLAAVLLQNRRLDLTLAE RGCTCLFGECCYVQNSGIVTEKVEIRDIQRRAEELRNTGFWGLLSOWMPWILP FLGLAAIILLFLPGPCTFNLLVNFVSRIEAVKLQMEPKVQSKTIYRRLDRPASP RSDVNDIKGTPPEISAAQPLLRNSAGSS"

1705. .1720

misc_feature

/note="splice acceptor site"

1959. .2694

LTR

ORIGIN

Query Match 90.1%; Score 1333.8; DB 9; Length 2694;

Best Local Similarity 93.8%; Pred. No. 0;

Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATACACTTTTCTTTACTTGTCTTACCCTTTCCCTCTCTACT 60

Db 284 ATGGCCCTCCCTTATCATATTTTCTTTACTTGTCTTTTACCCTCTTCTACTCTACT 343

Qy 61 GCACCCCTCCATGCTGCTGTACACCACTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120

Db 344 GCACCCCTCCATGCTGCTGTATGACCACTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 403

Qy 121 ACGGCGCTTCTTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGAAACTCC 180

Db 404 ATGACGCGTCCCGGAAATATTGATGCCCATCATGATAGAGTCTTCTAAGGAAACTCC 463

Qy 181 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATACTTGCCTACTCTTTGCATG 240

Db 464 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATACTTGCCTACTCTTTGCATG 523

Qy 241 CATGCAATATCTATTATTGGACAGGAAATGATTAATCTTAGTTGCTCTGAGGACTT 300

Db 524 CATGCAATATCTATTATTGGACAGGAAATGATTAATCTTAGTTGCTCTGAGGACTT 583

Qy 301 GGAGCCACTGCTGTGGACTTACTTACCCATACAGTATGCTGATGGGGTGAAT 360

Db 584 GGAGTCACTGCTGTGGACTTACTTACCCAACTGGTATGCTGATGGGGTGGAGTT 643

Qy 361 CAAGGTCAAGCAAGGAAACAAAGTAAAGAGCAATCTCCCACTGACCCGGGACAT 420

Db 644 CAAGATCAGCAAGGAAACAAAGTAAAGAGTATCTCCCACTACCCGGGTACAT 703

Qy 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTCATGAAACCTTCCGTACC 480

Db 704 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTCATGAAACCTTCCGTACC 763

Qy 481 CATACTCGCTGTGAGCTTATTAATACACCTCACTCGGTCCATGAGGTCTCAGCC 540

Db 764 CATACTCGCTGTGAGCTTATTAATACACCTCACTCGGTCCATGAGGTCTCAGCC 823

Qy 541 CAAACCTTACTAACTGTTGGATGTCCTCCCTGCACCTCAGGGCATACTTTCAATC 600

Db 824 CAAACCTTACTAACTGTTGGATGTCCTCCCTGAACTTTCAGGGCATATGTTCAATC 883

Qy 601 CTTGTTCTCTGAACAATGGAACAACTTTCAGCACAGAAATAAACACCACTTCCGTTTAGTA 660

Db 884 CTTGTTCTCTGAACAATGGAACAACTTTCAGCACAGAAATAAACACCACTTCCGTTTAGTA 943

Qy 661 GGAACCTTGTGTTTCAATCTGGAATAACCATACCTCAAACTCAGCTGTGTAATTTT 720

Db 944 GGAACCTTGTGTTTCAATCTGGAATAACCATACCTCAAACTCAGCTGTGTAATTTT 1003

Qy 721 AGCAATACTATAGACAAACCACTCCCAATGATCAGGTGGGTAAACCTCCCAACACGA 780

Db 1004 AGCAATACTATAGACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCCAACAA 1063

Qy 781 ATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTAACCTCAGCTATCATTTGTTG 840

Db 1064 ATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTAACCTCAGCTATCATTTGTTG 1123

Qy 841 AATGGCTCTTCAAGATCTATGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900

Db 1124 AATGGCTCTTCAAGATCTATGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183

Qy 901 ACTGAACAGATTTATACATCATGCTGCTTAAAGCCCAACAAAGAGTACCCATT 960

Db 1184 ACTGAACAGATTTATACATCATGCTGCTTAAAGCCCAACAAAGAGTACCCATT 1243

Qy 961 CTTCTCTTTTCTTATCAGAGCAGGAGTCTAGGAGACTAGGTACTGGCATTGGCAGTATC 1020

Db 1244 CTTCTCTTTTCTTATAGGAGCAGGAGTCTAGGTAGTACTGGCATTGGCAGTATC 1303

Qy 1021 ACAACCTTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAAACAG 1080

Db 1304 ACAACCTTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAAACAG 1363

Qy	1081	GTCAC	TGACTC	CTCGT	CACCTT	GCAAG	ATGA	CTTA	CTCCT	AGCAG	CAGT	AGTCC	CTT	1140						
Db	1364	GTCGC	CGACTC	CGCTG	TCACT	CGCA	AGAT	CAAC	TTAA	CTCCT	AGCAG	CAGT	AGTCC	CTT	1423					
Qy	1141	CAAA	ATCGA	AGAG	CTTT	AGAC	TTG	CTA	ACCG	CAAA	AGAG	GGGG	AACT	GTGTTT	ATTTTA	1200				
Db	1424	CAAA	ATCGA	AGAG	CTTT	AGAC	TTG	CTA	ACCG	CAAA	AGAG	GGGG	AACT	GTGTTT	ATTTTA	1483				
Qy	1201	GGAG	AAGAA	CGCT	GTAT	TAT	TAT	TAT	CA	AT	CC	AGAA	TTG	CTA	CTG	AGAA	AGTT	AAAGAA	1260	
Db	1484	GGGA	AGAA	TGCT	GTAT	TAT	TAT	TAT	CA	AT	CC	GAAT	CGT	CAC	TG	AGAA	AGTT	AAAGAA	1543	
Qy	1261	ATT	CGAG	AT	CGA	AT	CA	AT	CA	AT	CA	AT	CA	AT	CA	AT	CA	AT	CA	1320
Db	1544	ATT	CGAG	AT	CGA	AT	CA	AT	CA	AT	CA	AT	CA	AT	CA	AT	CA	AT	CA	1603
Qy	1321	CTC	AGC	CA	AT	GG	AT	GG	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	1380
Db	1604	CTC	AGC	CA	AT	GG	AT	GG	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	1663
Qy	1381	TTA	CTC	CT	CT	TT	GG	AC	CT	GT	AT	CT	TT	AA	CT	CT	CT	CT	CT	1440
Db	1664	CTA	CT	CT	CT	TT	GG	AC	CT	GT	AT	CT	TT	AA	CT	CT	CT	CT	CT	1723
Qy	1441	GA	AG	CT	GT	AA	AG	CT	CA	AG	AT	GG	CT	CT	CT	CT	CT	CT	CT	1481
Db	1724	GA	AG	CT	GT	AA	AG	CT	CA	AG	AT	GG	CT	CT	CT	CT	CT	CT	CT	1764

Search completed: February 21, 2005, 11:52:28
Job time : 6367.91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:25:25 ; Search time 763.929 Seconds
(without alignments)
11476.370 Million cell updates/sec

Title: US-09-319-156B-9
Perfect score: 1481
Sequence: 1 atggccctccattatcatac.....gtttacaaatggaacccca 1481

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1980s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1481	100.0	1481	2 AAV43217	AAV43217 Multiple
2	1481	100.0	1481	2 AAX29703	Aax29703 Clone C15
3	1481	100.0	1481	12 ADG14846	Adg14846 MSRV asso
4	1481	100.0	2030	3 AAG63826	Aag63826 Nucleotid
5	1461.8	98.7	1629	3 AAA96625	Aaa96625 DNA encod
6	1450	97.9	1481	9 ADB84400	Adb84400 MSRV-1 as
7	1389.8	93.8	2074	6 AAD41225	Aad41225 Human EMB
8	1335.4	90.2	2930	6 AAD24195	Aad24195 Human syn
9	1335.4	90.2	2946	2 AAX77526	Aax77526 Human sec
10	1335.4	90.2	2946	3 AAZ59468	Aaz59468 Human sec
11	1335.4	90.2	2946	10 ADC38776	Adc38776 Human CDN
12	1333.8	90.1	1617	5 AAH20070	Aah20070 HERV-W en
13	1333.8	90.1	2055	3 ABN97948	Abn97927 Human ret
14	1333.8	90.1	2599	3 ABN97927	Abn97927 Human ret
15	1333.8	90.1	8523	7 ADS30988	Ads30988 Human gen
16	1333.8	90.1	10499	3 ABN97929	Abn97929 Human ret
17	1333.8	90.1	56093	6 ABL61744	Abi61744 Colon ade
18	1330.6	89.8	2781	5 AAF55630	Aaf55630 Nucleotid
19	1324.2	89.4	7582	3 AAX59215	Aax59215 Human end
20	1322.6	89.3	7582	2 AAX25665	Aax25665 Complete

21	1308.2	88.3	2782	5 AAH20069	Aah20069 HERV-W en
22	1306.6	88.2	2782	2 AAX25661	Aax25661 Human end
23	1306.6	88.2	2782	3 AAX59211	Aax59211 5' non co
24	1298.2	87.7	3464	10 ADE09587	Ado09587 Novel DNA
25	1298.2	87.7	9502	10 ADF59718	Adf59718 Human con
26	1282	86.6	46340	3 ABN97978	Abn97978 Human ret
27	1278.8	86.3	1799	3 ABN97931	Abn97931 Human ret
28	1277.4	86.3	161334	11 ACN44334	Acn44334 Human gen
29	1239	83.7	2784	3 ABN97930	Abn97930 Human ret
30	1131.6	76.4	1894	4 AAI14608	Aai14608 Probe #45
31	1131.6	76.4	1894	4 ABA56337	Abas6337 Human foe
32	1131.6	76.4	1894	4 AAI35980	Aai35980 Probe #46
33	1131.6	76.4	1894	4 ABA45822	Abas4822 Human bre
34	1131.6	76.4	1894	4 ABA45822	Abas5978 Probe #44
35	1131.6	76.4	1894	4 ABA25978	Abas25978 Human bon
36	1131.6	76.4	1894	4 AAK04516	Aak04516 Human bra
37	1131.6	76.4	1894	4 ABS29670	Abs29670 Human liv
38	1131.6	76.4	1894	5 AAI04422	Aai04422 Probe #44
39	1131.6	76.4	1894	6 ABS04589	Abs04589 Human gen
40	1116.4	75.4	1948	2 AAX25659	Aax25659 Human end
41	1116.4	75.4	1948	3 AAS59209	Aas59209 5' non co
42	1048	70.8	6394	5 AAS84210	Aas84210 DNA encod
43	716.6	48.4	792	4 AAI23803	Aai23803 Probe #13
44	716.6	48.4	792	4 ABA68919	Abas68919 Human foe
45	716.6	48.4	792	4 AAI49115	Aai49115 Probe #17

ALIGNMENTS

RESULT 1
AAV43217
ID AAV43217 standard; cDNA; 1481 BP.

XX AAV43217;

DT 29-DEC-1998 (first entry)

XX Multiple sclerosis associated retrovirus fragment 5.

XX Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW env gene; rheumatoid arthritis-associated virus; ss.

XX Multiple sclerosis associated retrovirus.

Key Location/Qualifiers
CDS 1..1479

FT /*tag= a
FT /product= "Encodes protein AAV71068"
FT /transl_except= (pos:115-117, appears to code for a stop
FT codon)
FT /note= "CDS does not contain a stop codon"

XX WO9823755-A1.

PN 04-JUN-1998.

XX 26-NOV-1997; 97WO-IB001452.

XX 26-NOV-1996; 96US-00756429.

XX (INMR) BIO MERIEUX.

XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
PI WPI; 1998-322732/28.

XX P-PSDB; AAW71068.

XX New nucleic acid from retroviruses - useful for diagnosis, prevention and
FT treatment of, e.g. multiple sclerosis.

XX Disclosure; Page 184-185; 286pp; English.

XX The present sequence represents a multiple sclerosis (MS) associated
 CC retrovirus (MSRV) genomic fragment used in the method of the invention.
 CC The invention provides complete or partial genomic sequences of the MSRV-
 CC 1 pol gene, gag gene and env gene, and polypeptides encoded by these
 CC genes. The invention also provides antibodies raised against the
 CC polypeptides. The genomic sequences, polypeptides and antibodies are also
 CC claimed useful for diagnosing infection by MS and rheumatoid arthritis-
 CC associated viruses, and also for prevention and treatment of infection
 CC with these viruses
 XX
 SQ Sequence 1481 BP; 412 A; 410 C; 261 G; 398 T; 0 U; 0 Other;

Query Match 100.0%; Score 1481; DB 2; Length 1481;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTCTCTTACTGTCTTCTACCCCTTTTCGCTCTCACT 60
 DB 1 ATGGCCCTCCCTTATCATACTTTCTCTTACTGTCTTCTACCCCTTTTCGCTCTCACT 60
 QY 61 GCACCCCTCCATGCTGTGTACACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
 DB 61 GCACCCCTCCATGCTGTGTACACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
 QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC 180
 DB 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC 180
 QY 181 ACCTTCACCTGCCACACACCATATGCCCGCACTGTATACTCTGCCACTCTTTGCAATG 240
 DB 181 ACCTTCACCTGCCACACCATATGCCCGCACTGTATACTCTGCCACTCTTTGCAATG 240
 QY 241 CATGCAAACTACTTATTTGGACAGGAAATGATTAATCTAGTTGTCTGGAGGACTT 300
 DB 241 CATGCAAACTACTTATTTGGACAGGAAATGATTAATCTAGTTGTCTGGAGGACTT 300
 QY 301 GGAGCCACTGTCTGTTGGACTTACTTACCCATACAGTATGTCTGATGGGGGTGGAATT 360
 DB 301 GGAGCCACTGTCTGTTGGACTTACTTACCCATACAGTATGTCTGATGGGGGTGGAATT 360
 QY 361 CAAAGTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAT 420
 DB 361 CAAAGTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAT 420
 QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 DB 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 QY 481 CATACTCGGCTGGTGGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
 DB 481 CATACTCGGCTGGTGGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
 QY 541 CAAACCCCTACTAACTGTTGGATGTGCCTCCCTGCACCTCAGGCGCATACATTTCAATC 600
 DB 541 CAAACCCCTACTAACTGTTGGATGTGCCTCCCTGCACCTCAGGCGCATACATTTCAATC 600
 QY 601 CTGTTTCTGAAACAATGGAACAATTTTACAGACAGAAATAAACAACCACTTCCGTTTATGA 660
 DB 601 CTGTTTCTGAAACAATGGAACAATTTTACAGACAGAAATAAACAACCACTTCCGTTTATGA 660
 QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCTACCTCAAACTGCTGTAATTTT 720
 DB 661 GGACCTCTGTTTCCAACTGGAATAAACCCTACCTCAAACTGCTGTAATTTT 720
 QY 721 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
 DB 721 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
 QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCATCATGTTTGT 840
 DB 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCATCATGTTTGT 840

QY 841 AATGCTCTTCAGAACTATATGCTCTCTCTCATTTCTTAGTGCCCTTATGACCATCTAC 900
 DB 841 AATGCTCTTCAGAACTATATGCTCTCTCTCATTTCTTAGTGCCCTTATGACCATCTAC 900
 QY 901 ACTGAACAAGATTATACAACTCATGTCTAGTCTTAAAGCCCAACAACAAGAGTAGCCATT 960
 DB 901 ACTGAACAAGATTATACAACTCATGTCTAGTCTTAAAGCCCAACAACAAGAGTAGCCATT 960
 QY 961 CTTCTTTTGTATCAGACAGGAGTCTAGGAGACTAGGTACTGGCATTTGGCAGTAGTC 1020
 DB 961 CTTCTTTTGTATCAGACAGGAGTCTAGGAGACTAGGTACTGGCATTTGGCAGTAGTC 1020
 QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
 DB 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
 QY 1081 GTCACTGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCTAGCAGCAGTAGTCCTT 1140
 DB 1081 GTCACTGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCTAGCAGCAGTAGTCCTT 1140
 QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTGTTTATTTT 1200
 DB 1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTGTTTATTTT 1200
 QY 1201 GGAGAAGAACCTGTTTATTTATTAATCAATCCAGAAATTTGCTCACTGAGAAAGTTAAAGAA 1260
 DB 1201 GGAGAAGAACCTGTTTATTTATTAATCAATCCAGAAATTTGCTCACTGAGAAAGTTAAAGAA 1260
 QY 1261 ATTGAGATCGAATACAAATGATGAGCAGAGGCTTCAAAACACCGAAGCTGGGGCTC 1320
 DB 1261 ATTGAGATCGAATACAAATGATGAGCAGAGGCTTCAAAACACCGAAGCTGGGGCTC 1320
 QY 1321 CTCGACCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
 DB 1321 CTCGACCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
 QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 1440
 DB 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 1440
 QY 1441 GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481
 DB 1441 GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481

RESULT 2
 AAX29703
 ID AAX29703 standard; DNA; 1481 BP.
 XX
 AC AAX29703;
 XX
 DT 17-OCT-2003 (revised)
 DT 27-AUG-2003 (revised)
 DT 08-JUN-1999 (first entry)
 XX
 DE Clone C15 from MSRV-1.
 XX
 KW Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
 KW rheumatoid polyarthritis; ss.
 XX
 OS Viruses.
 XX
 PN FR2765588-A1.
 XX
 PD 08-JAN-1999.
 XX
 PF 07-JUL-1997; 97FR-00008816.
 XX
 PR 07-JUL-1997; 97FR-00008816.
 XX
 PA (INMR) BIO MERIEUX.
 XX
 DR WPI; 1999-098275/09.

PI Perron H, Beseme F, Bedin F, Paranhos-Baccalla G;
 PI Komurian-Pradel F, Jollivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
 XX WPI; 2004-032461/03.
 XX New isolated nucleic acid and their fragments having the pol gene of a
 PT retrovirus, useful for diagnosing, preventing and/or treating multiple
 PT sclerosis and/or rheumatoid arthritis.
 XX Disclosure; SEQ ID NO 105; 193pp; English.
 XX The invention relates to an isolated nucleic acid which comprises the pol
 CC gene of a retrovirus associated with multiple sclerosis or rheumatoid
 CC arthritis. The methods and compositions of the present invention are
 CC useful for diagnosing, preventing and/or treating multiple sclerosis
 CC and/or rheumatoid arthritis. The present sequence is used in the
 CC exemplification of the invention.
 XX Sequence 1481 BP; 412 A; 410 C; 261 G; 398 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 1481; DB 12; Length 1481;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTATCATACATCTTTCTCTTTACTGTTCTCTTACCCCTTTCCGCTCTCACT 60
 DB 1 ATGGCCCTCCCTATCATACATCTTTCTCTTTACTGTTCTCTTACCCCTTTCCGCTCTCACT 60
 QY 61 GCACCCCTCCATGCTGCTGTACACACAGTAGTCTCCCTTTACCAAGAGTTTCTATGAAGA 120
 DB 61 GCACCCCTCCATGCTGCTGTACACACAGTAGTCTCCCTTTACCAAGAGTTTCTATGAAGA 120
 QY 121 ACGGCGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAATCTCC 180
 DB 121 ACGGCGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAATCTCC 180
 QY 181 ACCTTCACCTGCCACACCCATATGCGCCCACTGCTATACTCTGCCACTCTTTGCATG 240
 DB 181 ACCTTCACCTGCCACACCCATATGCGCCCACTGCTATACTCTGCCACTCTTTGCATG 240
 QY 241 CATGCAAACTACTCATTTATTTGGAAGGAAATGATTAACTCTAGTTGCTCTGGAGGACTT 300
 DB 241 CATGCAAACTACTCATTTATTTGGAAGGAAATGATTAACTCTAGTTGCTCTGGAGGACTT 300
 QY 301 GGAGCCACTGCTGTTGGACTTACTTCAACCATACAGGATATGCTGATGGGGTGGAAAT 360
 DB 301 GGAGCCACTGCTGTTGGACTTACTTCAACCATACAGGATATGCTGATGGGGTGGAAAT 360
 QY 361 CRAAGTCAGCAAGAGAAACCAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
 DB 361 CRAAGTCAGCAAGAGAAACCAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
 QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 DB 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 QY 481 CATACTGCGCTGCTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
 DB 481 CATACTGCGCTGCTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
 QY 541 CAAACCCCTACTAATCTGATGTGCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
 DB 541 CAAACCCCTACTAATCTGATGTGCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
 QY 601 CCTGTTCTGAAACAAATGGAACAACTTCAGCACAGAAATAAACCACTTCGTTTGTAGTA 660
 DB 601 CCTGTTCTGAAACAAATGGAACAACTTCAGCACAGAAATAAACCACTTCGTTTGTAGTA 660
 QY 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAACCTCTGTTGTAATTT 720
 DB 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAACCTCTGTTGTAATTT 720
 QY 721 AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCACACGA 780

DB 721 AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCACACGA 780
 QY 781 ATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTTACCTCAGCTATCATTTGTTG 840
 DB 781 ATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTTACCTCAGCTATCATTTGTTG 840
 QY 841 AATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
 DB 841 AATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
 QY 901 ACTGAACAAGATTATTAACAATCATGTCTTAAGCCCAACAAACAAAGAGTACCCATT 960
 DB 901 ACTGAACAAGATTATTAACAATCATGTCTTAAGCCCAACAAACAAAGAGTACCCATT 960
 QY 961 CTTCCCTTTTGTATCAGAGCAGGAGTCTAGGAGAGTCTAGTCTGCGCATTTGGCAGATC 1020
 DB 961 CTTCCCTTTTGTATCAGAGCAGGAGTCTAGGAGAGTCTAGTCTGCGCATTTGGCAGATC 1020
 QY 1021 ACAACCTCTACTCAGTTTCTACTACAAATCTCTCAAGAAATAAATGTTGATCGAATGAAACAG 1080
 DB 1021 ACAACCTCTACTCAGTTTCTACTACAAATCTCTCAAGAAATAAATGTTGATCGAATGAAACAG 1080
 QY 1081 GTCACTGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTT 1140
 DB 1081 GTCACTGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTT 1140
 QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCCCAAGAGAGGGGAACTCTGTTATTTTAA 1200
 DB 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCCCAAGAGAGGGGAACTCTGTTATTTTAA 1200
 QY 1201 GGAGGAAGAACCTGTTATTATTTAATCAATCCAGAAATTTGTCACCTGAGAAAGTTAAAGAA 1260
 DB 1201 GGAGGAAGAACCTGTTATTATTTAATCAATCCAGAAATTTGTCACCTGAGAAAGTTAAAGAA 1260
 QY 1261 ATTCCAGATCGAATACAAATGTAGCAGAGAGGCTTCAAAACACCGAAGCTGGGGCTC 1320
 DB 1261 ATTCCAGATCGAATACAAATGTAGCAGAGAGGCTTCAAAACACCGAAGCTGGGGCTC 1320
 QY 1321 CTCAGCAATGGATGCTGCTGCTCCCTTCTAGGACCTCTAGCAGCTCTAATATTG 1380
 DB 1321 CTCAGCAATGGATGCTGCTGCTCCCTTCTAGGACCTCTAGCAGCTCTAATATTG 1380
 QY 1381 TTACTCTCTTTGGACCTCTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAAT 1440
 DB 1381 TTACTCTCTTTGGACCTCTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAAT 1440
 QY 1441 GAAGCTGTAAAGCTACAGATGCTTCAAAATGGAACCCCA 1481
 DB 1441 GAAGCTGTAAAGCTACAGATGCTTCAAAATGGAACCCCA 1481
 RESULT 4
 AAA63826
 ID AAA63826 standard; DNA; 2030 BP.
 XX
 AC AAA63826;
 XX
 DT 06-AUG-2003 (revised)
 DT 04-DEC-2000 (first entry)
 XX
 DE Nucleotide sequence of the MSRV-1 3' env and LTR regions.
 XX MSRV-1; pol region; long terminal repeat; LTR; RU5 region; retrovirus;
 KW ss.
 XX Multiple sclerosis associated retrovirus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1629
 FT /tag= a
 FT /note= "Contains one termination codon"
 FT sig_peptide 1..81


```

FT CAAT_signal /*tag= b
FT 1800..1807
FT /*tag= c
FT CAAT_signal 1858..1864
FT /*tag= d
FT TATA_signal 1906..1911
FT /*tag= e
FT polyA_signal 1996..2002
FT /*tag= f
XX WO200047745-A1.
XX PN
XX PD
XX PR
XX PF 15-FEB-2000; 2000WO-1B000159.
XX PP
XX PR 15-FEB-1999; 99EP-00420041.
XX PP
XX PA (INMR ) BIO MERIEUX.
XX PI
XX PT Paranhos-Baccala G, Perron H, Komurian-Pradel P;
XX PR
XX DR WPI; 2000-506097/45.
XX DR P-PSDB; AAB08195.
XX XX
XX PT Nucleotide fragment of LTR-RUS region from Multiple Sclerosis retrovirus
XX PT (MSRV) used to detect the presence of MSRV-1 retrovirus in a biological
XX PT sample.
XX PS Disclosure; Fig 2; 23pp; English.
XX CC
XX CC The present sequence represents the nucleotide sequence corresponding to
XX CC the 3' env region and long terminal repeat sequences from clone Cl6 of
XX CC Multiple Sclerosis retrovirus (MSRV-1). The specification describes a
XX CC long terminal repeat (LTR)-RUS region which encodes the expression of a
XX CC MSRV-1 protein. This is unusual for LTRs, in particular in the RUS
XX CC region. The sequence includes CAAT and TATA signals which are present in
XX CC the U3 and R regions and are not directed towards the CDS indicated in
XX CC the features table. Probes and antibodies to the MSRV-1 retrovirus
XX CC protein and encoding polynucleotide sequences are used to detect the
XX CC presence of MSRV-1 retrovirus in a biological sample. (Updated on 06-AUG-
XX CC 2003 to correct OS field.)
XX XX
XX SQ Sequence 2030 BP; 574 A; 559 C; 387 G; 510 T; 0 U; 0 Other;
Query Match 100.0%; Score 1481; DB 3; Length 2030;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATACTTTTCTTCTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
DB 1 ATGGCCCTCCCTTATCATACTTTTCTTCTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
QY 61 GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
DB 61 GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
QY 121 ACGCGGCTTCTGGAAATATTTGATGCCCATCATATATAGGAGTTTATCTAAGGGAACTCC 180
DB 121 ACGCGGCTTCTGGAAATATTTGATGCCCATCATATATAGGAGTTTATCTAAGGGAACTCC 180
QY 181 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATTAACCTGCGCACTTTTGCATG 240
DB 181 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATTAACCTGCGCACTTTTGCATG 240
QY 241 CATGCAATATCTCATTTATTTGACAGGAAATATTAATCTTACTGTTCTTGGAGGACTT 300
DB 241 CATGCAATATCTCATTTATTTGACAGGAAATATTAATCTTACTGTTCTTGGAGGACTT 300
QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGATGTCCTGATGGGGTGGAAAT 360
DB 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGATGTCCTGATGGGGTGGAAAT 360

```

```

QY 361 CAAGTCTAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
DB 361 CAAGTCTAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
DB 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
QY 481 CATACTCGCCTGGTGAAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
DB 481 CATACTCGCCTGGTGAAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
QY 541 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTGCACTTCCAGGCCATAGATTTCAATC 600
DB 541 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTGCACTTCCAGGCCATAGATTTCAATC 600
QY 601 CCTGTTCTGTAACAATGGAACTTACGACAGAAATAAACACCACTTCCGTTTTAGTA 660
DB 601 CCTGTTCTGTAACAATGGAACTTACGACAGAAATAAACACCACTTCCGTTTTAGTA 660
QY 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTT 720
DB 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTT 720
QY 721 AGCAATATATAGACAAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
DB 721 AGCAATATATAGACAAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTGCTGTGCTACCTCAGCCTATCATTTGTTG 840
DB 781 ATAGTCTGCTACCTCAGGAATATTTTGTGCTGTGCTACCTCAGCCTATCATTTGTTG 840
QY 841 AATGGCTCTTCAAGATCTATGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
DB 841 AATGGCTCTTCAAGATCTATGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
QY 901 ACTGAAACAAGATTTATACAAATCATGTGCTGTAACCTTAAGCCCAACAAGAGTACCCATT 960
DB 901 ACTGAAACAAGATTTATACAAATCATGTGCTGTAACCTTAAGCCCAACAAGAGTACCCATT 960
QY 961 CTTCTCTTTTGTATCAGACAGGAGTCTAGGAGACTAGGTACTGCACTGTCAGTATC 1020
DB 961 CTTCTCTTTTGTATCAGACAGGAGTCTAGGAGACTAGGTACTGCACTGTCAGTATC 1020
QY 1021 ACAACCTCTACTCAGTCTTCTACTACAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
DB 1021 ACAACCTCTACTCAGTCTTCTACTACAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
QY 1081 GTCACCTGACTCCCTGTCACCTTTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1140
DB 1081 GTCACCTGACTCCCTGTCACCTTTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1140
QY 1141 CAAAATCGAAGAGCTTTAGACTTTGCTAACCGCAAGAGGGGAACTGTTTATTTTAA 1200
DB 1141 CAAAATCGAAGAGCTTTAGACTTTGCTAACCGCAAGAGGGGAACTGTTTATTTTAA 1200
QY 1201 GGAGAGAAACGCTGTTTATTTATGTTTAACTCAGAAATTTCTCACTGAGAAAGTTTAAAGAA 1260
DB 1201 GGAGAGAAACGCTGTTTATTTATGTTTAACTCAGAAATTTCTCACTGAGAAAGTTTAAAGAA 1260
QY 1261 ATTCTGAGATCGAATACAAATGTAGACAGAGAGCTTCAAAACACCGCAACGCTGGGGCTC 1320
DB 1261 ATTCTGAGATCGAATACAAATGTAGACAGAGAGCTTCAAAACACCGCAACGCTGGGGCTC 1320
QY 1321 CTGAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1380
DB 1321 CTGAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1380
QY 1381 TTAATCTCTTTTGAGCCCTGTATCTTTAACTCTCTTCTTAAAGTTTCTCTTCCAGAAAT 1440
DB 1381 TTAATCTCTTTTGAGCCCTGTATCTTTAACTCTCTTCTTAAAGTTTCTCTTCCAGAAAT 1440
QY 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA 1481

```


DB 1201 GGAGAGACGCTGTTATATGTTATCAATCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
 QY 1261 ATTGAGATCGAATACATAGTAGAGAGAGAGCTTCAAAACACCGAAGCTGGGCGCTC 1320
 DB 1261 ATTGAGATCGAATACATAGTAGAGAGAGAGCTTCAAAACACCGAAGCTGGGCGCTC 1320
 QY 1321 CTCAGCAATGGATGCCCTGGGTTCTCCCTCTTCTAGGACCTCTAGCAGCTCTAAATATG 1380
 DB 1321 CTCAGCAATGGATGCCCTGGGTTCTCCCTCTTCTAGGACCTCTAGCAGCTCTAAATATG 1380
 QY 1381 TTACTCTCTTTGGACCTCTGATCTTTAACTCTTTGTTAAGTTCTCTTCCGAAAT 1440
 DB 1381 TTACTCTCTTTGGACCTCTGATCTTTCAACTCTCTTTGTTAAGTTCTCTTCCGAAAT 1440
 QY 1441 GAAGCTGTAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481
 DB 1441 GAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCCA 1481

RESULT 6
 ADB84400
 ID ADB84400 standard; DNA; 1481 BP.

AC ADB84400;
 XX
 DT 04-DEC-2003 (first entry)
 XX

MSRV-1 associated DNA sequence #19.

DE ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
 KW reverse transcriptase; ribonuclease H.
 XX

OS Unidentified;.

XX US2003039664-A1.

XX 27-FEB-2003.

XX 26-NOV-1997; 97US-00979847.

XX 26-NOV-1996; 96US-00756429.

XX (PERR/) PERRON H.

XX (BESE/) BESEME F.

XX (BEDI/) BEDIN F.

XX (PARA/) PARANHOS-BACCALA G.

XX (KOMU/) KOMURIAN-PRADEL F.

XX (JOLI/) JOLIVET-REYNAUD C.

XX (MAND/) MANDRAND B.

XX (GARS/) GARSON J A.

XX (TUKE/) TUKE P W.

XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;

PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B,

XX WPI; 2003-512253/48.

XX New isolated or purified nucleic acid associated with multiple sclerosis

PT and/or rheumatoid arthritis, useful for detecting a virus associated with

PT multiple sclerosis or rheumatoid arthritis in a biological sample.

XX Claim 31; Page 77-78; 193pp; English.

XX The invention relates to an isolated or purified nucleic acid from a

XX virus associated with multiple sclerosis and/or rheumatoid arthritis,

CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise

CC primer for the amplification by polymerisation of a nucleic acid of a

CC viral material associated with multiple sclerosis or rheumatoid

CC arthritis, a polypeptide exhibiting an inhibitory activity on the

CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,

CC and an antibody directed against the MSRV-1 virus obtained by

CC immunologically reacting a human or animal body or cells with an

CC immunogenic agent consisting of the antigenic polypeptide defined above.

CC The nucleic acids are useful for detecting a biological sample a virus

CC associated with multiple sclerosis or rheumatoid arthritis, or for

CC detecting in a biological sample, the presence of or exposure to a virus

CC associated with multiple sclerosis or rheumatoid arthritis. The present

CC sequence is a claimed MSRV-associated sequence whose identity cannot be

CC accurately determined. Note: The SEQ ID numbers for the sequences as

CC displayed in the main body of the patent do not match the SEQ ID numbers

CC in the sequence listing. Consequently those sequences mentioned in the

CC claims may not be the sequences the authors intended to claim.

XX

XX SQ Sequence 1481 BP; 405 A; 403 C; 257 G; 385 T; 0 U; 31 Other;

Query Match 97.9%; Score 1450; DB 9; Length 1481;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1450; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACACTTTTCTTCTTACTGTTCTTACCCCTTCGCTCTCACT 60

DB 1 ATGGCCCTCCCTTATCATACACTTTTCTTCTTACTGTTCTTACCCCTTCGCTCTCACT 60

QY 61 GCACCCCTCCATGCTGTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120

DB 61 GCACCCCTCCATGCTGTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120

QY 121 ACGCGCTTCTCGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180

DB 121 ACGCGCTTCTCGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCN 180

QY 181 ACCTTCAGTCCCAACCCATATATGATGCCCACTGCTATTAACCTCTGCGACTCTTCGATC 240

DB 181 ACCTTCAGTCCCAACCCATATATGATGCCCACTGCTATTAACCTCTGCGACTCTTCGATC 240

QY 241 CATGCAATTAATCTATTGAGAGGAGAAATGATTAACTCTAGTTGTCTGAGGAGACTT 300

DB 241 CATGCAATTAATCTATTGAGAGGAGAAATGATTAACTCTAGTTGTCTGAGGAGACTN 300

QY 301 GGAGCCACTGTCTGTTGGACTTTACTTACCCTACCAATACAGTATGTCTGATGGGTTGAAAT 360

DB 301 GGAGCCACTGTCTGTTGGACTTTACTTACCCTACCAATACAGTATGTCTGATGGGTTGAAAT 360

QY 361 CAAGGTGAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCCACTGACCCCGGGACAT 420

DB 361 CAAGGTGAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCCACTGACCCCGGGACAN 420

QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCCCGTACC 480

DB 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCCCGTACN 480

QY 481 CATACTCGCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540

DB 481 CATACTCGCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCN 540

QY 541 CAAAACCTTAACTGTTGGATGTGCTTCCCTCGACTTCAGGGCATACATTTCAATC 600

DB 541 CAAAACCTTAACTGTTGGATGTGCTTCCCTCGACTTCAGGGCATACATTTCAATN 600

QY 601 CCTGTTCTTGAACAATGGAACTTTCAGCAGAGAAATAAACACCACTTCCTGTTTAGTA 660

DB 601 CCTGTTCTTGAACAATGGAACTTTCAGCAGAGAAATAAACACCACTTCCTGTTTAGTN 660

QY 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTT 720

DB 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACTGTGTAAATTN 720

QY 721 AGCAATACTATAGACACACACCTCCCATGATGAGTGGGTAAACCTCCACACGA 780

Db	721	AGCA	TACTATAGACACAA	CCAGCTCCCAATG	CATCAGGTGGGTAA	CACCTCCCA	CACGN	780
Qy	781	ATAG	TGCGCTACCC	TCAGGAATATTTT	TGCTGTGGTACCT	CAGCCTCAT	CATTGTTG	840
Db	781	ATAG	TGCGCTACCC	TCAGGAATATTTT	TGCTGTGGTACCT	CAGCCTCAT	CATTGTTT	840
Qy	841	AATGG	CTCTTCAGAA	CTATGTGCTTCCT	CTCATTTT	TAGTGCCCT	TGACCACTAT	900
Db	841	AATGG	CTCTTCAGAA	CTATGTGCTTCCT	CTCATTTT	TAGTGCCCT	TGACCACTAT	900
Qy	901	ACTG	AAACAAGATT	TATACAATCAT	TGCTACCTAA	GCCCAACA	AAAGAGTACCCAT	960
Db	901	ACTG	AAACAAGATT	TATACAATCAT	TGCTACCTAA	GCCCAACA	AAAGAGTACCCAT	960
Qy	961	CTTTC	TTTTTGTAT	CAGACGAGG	GTCTAGG	CAGACTAG	GTACTGGCAT	1020
Db	961	CTTTC	TTTTTGTAT	CAGACGAGG	GTCTAGG	CAGACTAG	GTACTGGCAT	1020
Qy	1021	ACAAC	CTCTACTCAG	TTCTACTCAAA	CTATCTCA	AGAAATAA	TGGTGACATG	1080
Db	1021	ACAAC	CTCTACTCAG	TTCTACTCAAA	CTATCTCA	AGAAATAA	TGGTGACATG	1080
Qy	1081	GTCA	CTGACTCCCT	GTGTCACCTT	GCAAGATCA	ATTAACTCC	TAGCAGAGTAGTCC	1140
Db	1081	GTCA	CTGACTCCCT	GTGTCACCTT	GCAAGATCA	ATTAACTCC	TAGCAGAGTAGTCC	1140
Qy	1141	CAAA	ATCGAAGAGCT	TTTAGACTT	GTCTAA	CGCCAAAGAG	GGGGAACCTCTTTATTTT	1200
Db	1141	CAAA	ATCGAAGAGCT	TTTAGACTT	GTCTAA	CGCCAAAGAG	GGGGAACCTCTTTATTTT	1200
Qy	1201	GGAG	AAGAACGCTG	TATATGTTAA	TCAATCCAG	AATTTGCT	ACTGAGAAGTTAA	1260
Db	1201	GGAG	AAGAACGCTG	TATATGTTAA	TCAATCCAG	AATTTGCT	ACTGAGAAGTTAA	1260
Qy	1261	ATTC	GAGATCGAAT	ACAATG	TAGACGAGAG	GCTTCA	AAACCGCAACGCTGGG	1320
Db	1261	ATTC	GAGATCGAAT	ACAATG	TAGACGAGAG	GCTTCA	AAACCGCAACGCTGGG	1320
Qy	1321	CTCA	GCCAAATGG	ATGCCCTT	GGTTCCTCC	CTTTAG	GCCTCTAGCAGCTCTAA	1380
Db	1321	CTCA	GCCAAATGG	ATGCCCTT	GGTTCCTCC	CTTTAG	GCCTCTAGCAGCTCTAA	1380
Qy	1381	TTACT	CTCTTTTGG	ACCTGTATCTT	TAACTCTCT	TGTTAAG	TTTGTCTCTCCAG	1440
Db	1381	TTACT	CTCTTTTGG	ACCTGTATCTT	TAACTCTCT	TGTTAAG	TTTGTCTCTCCAG	1440
Qy	1441	GAAG	CTGTAAG	CTACAGATGG	CTTACA	AAATGGA	ACCCCA	1481
Db	1441	GAAG	CTGTAAG	CTACAGATGG	CTTACA	AAATGGA	ACCCCA	1481

RESULT 7	
AAD41225	
ID	AAD41225 standard; cDNA; 2074 BP.
XX	
XX	
AC	AAD41225;
XX	
XX	30-OCT-2002 (first entry)
DT	
XX	
DE	Human EMBRY-2 cDNA.
XX	
XX	Human; embryogenesis associated protein; AIDS; reproductive disorder;
KW	infertility; endometriosis; endometrial tumour; inflammatory disorder;
KW	autoimmune disorder; acquired immune deficiency syndrome; transgenic;
KW	ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;
KW	embry-2. allergen; gene therapy; gene; ss.
KW	

AA	Homo sapiens.	Location/Qualifiers
OS		64..1692
XX		/*tag= a
FX		/product= "EMBRY-2 protein"
FT		
FT		
FT		

FT	sig_peptide	64..123
FT		/tag= b
FT	mat_peptide	124..1689
FT		/tag= c
FT		/product= "Mature EMBRY-2 protein"
XX		
XX	W0200248362-A2.	
PN		
XX		
PD	20-JUN-2002.	
XX		
XX	14-NOV-2001; 2001WO-US043956.	
PF		
XX	15-NOV-2000; 2000US-0249407P.	
PR		
XX	(INCY-) INCYTE GENOMICS INC.	
PA		
XX		
PI	Ramkumar J, Arvizu C;	
XX		
DR	WPI; 2002-537629/57.	
DR	P-PSDB; AAE25054.	
XX		
PT	New polypeptides of human embryogenesis associated proteins for screening	
PT	modulators useful for treating or preventing disorders e.g.	
PT	endometriosis, infertility, allergy, pre-eclampsia.	
XX		
PS	Claim 59; Page 96-97; 97pp; English.	
XX		
CC	The invention relates to human embryogenesis associated proteins (EMBRYP)	
CC	and nucleic acid molecules encoding such proteins. EMBRY sequences are	
CC	useful for screening modulators useful for treating or preventing	
CC	disorders associated with abnormal expression of EMBRY. The disorders	
CC	treated include reproductive disorders such as infertility,	
CC	endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory	
CC	disorder such as acquired immune deficiency syndrome (AIDS), allergies,	
CC	contact dermatitis; disorders of the placenta such as pre-eclampsia,	
CC	abruptio placentae etc. Sequences of the invention are also useful for	
CC	analysing a proteome of a tissue or a cell type. EMBRY proteins are	
CC	useful as immunogens for preparing antibodies. Polynucleotides of the	
CC	invention are useful for creating knockin humanised animals or transgenic	
CC	animals to model human diseases. They are also used in gene therapy. The	
CC	present sequence is human EMBRY-2 cDNA	
XX		
SQ	Sequence 2074 BP; 583 A; 567 C; 392 G; 531 T; 0 U; 1 Other;	
Query Match 93.8%; Score 1389.8; DB 6; Length 2074;		
Best Local Similarity 96.2%; Pred. No. 0;		
Matches 1424; Conservative 0; Mismatches 57; Indels 0; Gaps 0		
QY	1	ATGGCCCTCCCTTATCATACATTTTCTCTTTACTGTGTCCTTACCCCCTTCGCCTCACT 60
Dd	64	ATGGCCCTCCCTTATGTAATTTTTCTTTACTGTTCTCTCACACCCTTCAGTCTCACT 123
QY	61	GCACCCCTCCATGCTGCTGTACAAACAGTAGTCCCCCTTACCAAGAGTTTCTATGAAGA 120
Dd	124	GCACCCCTCCTCATGCCACTGTAGGACCAGTAGTCCCCCTTACCAAGCGTTTCTATGGAGA 183
QY	121	ACGCGGCTCTCGGAAATATTGATGCCCCCATCATATAGSAGTTTATCTAAGGAAATCC 180
Dd	184	ATGCGGCGTGCCAGACATAATTGATGCCCCCATCGTATAGAGTTTATCTAAGGAAATCC 243
QY	181	ACCTTTCAC TGCCCAACCCATATGCCCGCACTGCTATAACTCTGCCACTTTTTCGATG 240
Dd	244	GCTTTCACGCGCCACCCATATGCCCACTGCTATAACTCTGCCACTTTTGTATG 303
QY	241	CATGCAAAATCATATTATGACAGGGAATATGATTATCTAGTTGTCCTGGAGACTT 300
Dd	304	CATGCAAAATCATATTATGACAGGGAATATGATTATCTAGTTGTCCTGGAGACTT 363
QY	301	GGAGCCACTGCTGTTTGACTTACTTTCACCCATACCACTATGTCCTGATGGGGTGGAAAT 360
Dd	364	GGAGCCACTATCTGTTTGACTTACTTTCACCCATACCGGTAATGTCCTGATGGGGTGGAGTT 423
QY	361	CAAAGTCAGGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCCACTGACCCGGGACAT 420

Db 1050 ATGAGCGTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTCTAAGGAACCCCC 1109
Qy 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCOACTCTTTGCAATG 240
Db 1110 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCOACTCTTTGCAATG 1169
Qy 241 CATGCAATACTCATTTATGGACAGGAAATGATTAATCCCTAGTTGCTCGAGACATT 300
Db 1170 CATGCAATACTCATTTATGGACAGGAAATGATTAATCCCTAGTTGCTCGAGACATT 1229
Qy 301 GGAGCCACTGTCTGTGGACTTACTTACCCATACCAAGTATGCTGATGGGGGTGGAATT 360
Db 1230 GGAGTCACTGTCTGTGGACTTACTTACCCAACTGGTATGCTGATGGGGGTGGAATT 1289
Qy 361 CAAAGTCAAGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
Db 1290 CAAAGTCAAGCAAGAGAAAAAACAAGTAAAGGAAGTAACTCCCAACTCAACCCGGGTACAT 1349
Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db 1350 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC 1409
Qy 481 CATACTCGCTGTGAGCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 1410 CATACTCGCTGTGAGCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 1469
Qy 541 CAAACCCCTACTAATCTGTGGATGTCCTCCCTCGCACTTCAGGCAATACATTTCAATC 600
Db 1470 CAAACCCCTACTAATCTGTGGATGTCCTCCCTCGCACTTCAGGCAATATGTTCAATC 1529
Qy 601 CCGTGTCTGMAAATGGACAACTTCAGCAGAGAATAAACAACCACTTCGGTTTTAGTA 660
Db 1530 CCGTGTACCTGAAACAATGGACAACTTCAGCAGAGAATAAACAACCACTTCGGTTTTAGTA 1589
Qy 661 GGACCTCTGTGTTTCCAACTCTGAAATAAACCCATACCTCAAACTCACTGTGTAAATTT 720
Db 1590 GGACCTCTGTGTTTCCAACTCTGAAATAAACCCATACCTCAAACTCTGTGTAAATTT 1649
Qy 721 AGCAATACTATAGACAAACCCAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
Db 1650 AGCAATACTATAGACAAACCCAGCTCCCAATGATCAGGTGGGTAACTCCTCCACACAA 1709
Qy 781 ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
Db 1710 ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTG 1769
Qy 841 AATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCTATGACCAATCTAC 900
Db 1770 AATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCTATGACCAATCTAC 1829
Qy 901 ACTGAACAAGATTTATACATCATGTGCTACCTAAGCCCCCAACAAGAAGTACCCATT 960
Db 1830 ACTGAACAAGATTTATACAAATATGTATATCTTAAGCCCGGCAACAAAGAGTACCCATT 1889
Qy 961 CTTCTCTTTGTTATCAGACAGAGTGTAGGACAGACTAGGTACTGGCATTTGGCATATC 1020
Db 1890 CTTCTCTTTGTTATAGACAGAGTGTAGGACAGACTAGGTACTGGCATTTGGCGGTATC 1949
Qy 1021 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAATAAATGGTGACATGGAACAG 1080
Db 1950 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAATAAATGGGACATGGAACAG 2009
Qy 1081 GTCACGTACCTCCTGTGTCACCTTGCAGATCACTAACTCCCTACGACAGTAGTCCCTT 1140
Db 2010 GTCGCCGACTCCCTGTGTCACCTTGCAGATCACTAACTCCCTACGACAGTAGTCCCTT 2069
Qy 1141 CAAATCGAAGAGCTTTAGACTTGTAACTGCTAAACCCGCAAAAGAGGGGAACTGTTTAT 1200
Db 2070 CAAATCGAAGAGCTTTAGACTTGTAACTGCTAAACCCGCTGAAGAGGGGAACTGTTTAT 2129
Qy 1201 GGAGAGAAACGCTGTTATTTATGTTAATCAATCAGAAATTTGTCATGAGAAAGTTAAAGAA 1260
Db 2130 GGGGAAGAAATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 2189

Qy 1261 ATTGAGATCGAATACAAATGTAGAGCAGAGGCTTCAAAACACCGAACGCTGGGGCCTC 1320
Db 2190 ATTGAGATCGAATACAAACGTAGAGCAGAGGCTTCAAAACACTGGACCTGGGGCCTC 2249
Qy 1321 CTCAGCCAAATGGATGCCCTGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATG 1380
Db 2250 CTCAGCCAAATGGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATG 2309
Qy 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTAAAGTTTGTCTTCCAGAAAT 1440
Db 2310 CTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTAACTCTTGTAACTTTGTCTTCCAGAAATC 2369
Qy 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 2370 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCAA 2410

RESULT 9
AAX77526
ID AAX77526 standard; cDNA; 2946 BP.
XX
AC AAX77526;
XX
10-AUG-1999 (first entry)
XX
Human secreted protein AJ172_2 cDNA.
XX
Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW cell proliferation; cell differentiation; suppressor; tumour inhibition;
KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
KW cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.
XX
Homo sapiens.
OS
XX
WO9926972-A1.
XX
03-JUN-1999.
XX
17-NOV-1998; 98W-US024614.
XX
21-NOV-1997; 97US-00976110.
PR 18-MAY-1998; 98US-00080478.
PR 20-OCT-1998; 98US-00175928.
XX
(GEMY) GENETICS INST INC.
XX
Jacobs K, McCooy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M;
XX
WPI; 1999-357813/30.
DR P-PSDB; AAY08622.
XX
New polynucleotides encoding secreted proteins.
XX
Claim 13a; Page 100-101; 142pp; English.
XX
This invention describes novel human secreted proteins encoded by
CC polynucleotides isolated from human adult testes, adult brain, adult
CC blood or adult placenta, or murine adult bone marrow or thymus cDNA
CC libraries. The products of the invention are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity,
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful for

XX PS Claim 14; Page 107-108; 149pp; English.

CC This is the human secreted protein AJ172.2 nucleotide sequence, obtained from a human adult testes cDNA library. The invention relates to secreted human and murine proteins. The polynucleotides and proteins are predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals. Detection of the levels of the proteins can be used for the diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents which modulate the expression or function of the proteins may be used for treating a neoplastic disease and inhibiting metastasis. Other suggested activities include nutritional activity (e.g. in feeds), cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, and thrombolytic activity, chemotactic/chemokinetic activity, haemostatic and thrombotic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, and tumour inhibition activity. The polynucleotide sequences are also stated to be useful for gene therapy

XX SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

Query Match 90.2%; Score 1335.4; DB 3; Length 2946;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTATCATATCTTTCTTTTACTGTCTTACCTTACCCCTTTGCTCTCACT 60
DB 928 ATGGCCCTCCCTATCATATCTTTCTTTTACTGTCTTACCTTACCCCTTTTCACTCTCACT 987

QY 61 GCACCCCTCCATGCTGCTGACACACAGTACCTCCCTTACCAAGAGTTCTATGAAGA 120
DB 988 GCACCCCTCCATGCTGCTGATGACACAGTACCTCCCTTACCAAGAGTTCTATGGAGA 1047

QY 121 AGCGCGCTTCTCGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB 1048 ATGCAGCGTCCCGAAATATTGATGCCCATCATGATAGGAGTTCTTAAGGGAACCCCC 1107

QY 181 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATACTTGCCTCTTTGCATG 240
DB 1108 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATACTTGCCTCTTTGCATG 1167

QY 241 CATGCAATACCTATTATGGACGGAAGAAATGATTAATCCCTAGTTGCTCGAGAGCTT 300
DB 1168 CATGCAATACCTATTATGGACGGAAGAAATGATTAATCCCTAGTTGCTCGAGAGCTT 1227

QY 301 GGAGCCACTGTCTGTTGGACTTACTTTCACCCATACCCAGTATGCTGATGGGGTGGAAAT 360
DB 1228 GGAGTCACTGTCTGTTGGACTTACTTTCACCCAACTGGTATGCTGATGGGGTGGAGTT 1287

QY 361 CAAAGTCAAGCAAGAGAAAACAAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT 420
DB 1288 CAAAGTCAAGCAAGAGAAAACAAAGTAAAGAGCAATCTCCAACTGACCCGGGTACAT 1347

QY 421 AGCACCCTAGCCCTACAAAGAGCTAGTTCTCTCAAACTACATGAACCCCTCGGTACC 480
DB 1348 GGCACTCTAGCCCTTACAAAGAGCTAGATCTCTCAAACTACATGAACCCCTCGGTACC 1407

QY 481 CATACTCGGCTGGTGAAGCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
DB 1408 CATACTCGGCTGGTGAAGCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCGGC 1467

QY 541 CAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCATCTCAGGCCATACATTTCAATC 600
DB 1468 CAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCATCTCAGGCCATATGTTTCAATC 1527

QY 601 CCTGTTCTGAACTAAGCAACTTTCAGCAGAGAAATAACACCACTTCCGTTTTAGTA 660
DB 1528 CCTGTTACCTGAACATGGAACTTTCAGCAGAGAAATAACACCACTTCCGTTTTAGTA 1587

QY 661 GGACCTCTTGTGTTTCCAACTCTGAAATTAACCCATACCTCAAACTCTGTTGAAATTT 720

DB 1588 GGACCTCTTGTTCCTCAATCTCGAAATAACCCATACCTCAAACTCAGCTGTGTAAATTT 1647

QY 721 AGCAATACTATAGACAAACCCAGCTCCCAATGCAATCAGGTGGTAAACACCTCCCAACAGA 780

DB 1648 AGCAATACTATATACAAACCACTCCCAATGCAATCAGGTGGTAACTCTCTCCCAACAA 1707

QY 781 ATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCCTCATATTGTTG 840

DB 1708 ATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCCTCATGTTGTTG 1767

QY 841 AATGCTCTTCAGAAATCTATGCTTCTCTCATCTTTAGTGCCCTTATGACCATCTAC 900

DB 1768 AATGCTCTTCAGAAATCTATGCTTCTCTCATCTTTAGTGCCCTTATGACCATCTAC 1827

QY 901 ACTGAACAAGATTTATCAATCATGCTCTACCTAAGCCCAACAAAGAGTACCCATT 960

DB 1828 ACTGAACAAGATTTATCAATATGCTATATCTAAGCCCAACAAAGAGTACCCATT 1887

QY 961 CTTCTCTTTGTTATCAGAGCAGAGTGTCTAGCAGACTAGGTACTTGGCAATGGCAGTATC 1020

DB 1888 CTTCTCTTTGTTATAGGAGCAGAGTGTCTAGTGTGCTAGGTACTTGGCAATGGCAGTATC 1947

QY 1021 ACAACCTCTACTCAGTCTTACTACAAATATCTCAAGAAATAAATGGTGACATGGAACAG 1080

DB 1948 ACAACCTCTACTCAGTCTTACTACAAATATCTCAAGAACTAAATGGGACATGGAACGG 2007

QY 1081 GTCCTGACTCCCTGGTCCACCTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT 1140

DB 2008 GTGCGGACTCCCTGGTCCACCTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT 2067

QY 1141 CAAATCCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGGGAACCTGTTTATTTTAA 1200

DB 2068 CAAATCCGAAGAGCTTTAGACTTGTCTAACCCCTGAAGAGGGGGAACCTGTTTATTTTA 2127

QY 1201 GGAGAAAGACCTGTTTATTTATTCATCAATCCAGAAATGTCACAGAAAGTTAAAGAA 1260

DB 2128 GGGGAAGAAATCTGTTTATTTATTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 2187

QY 1261 ATTGAGATCAATACATATGATAGCAGAGAGCTTCAAAACACCAACGCTGGGGCCTC 1320

DB 2188 ATTGAGATCAATACATCAACGATAGCAGAGAGCTTCAAAACACCTGGACCTTGGGGCCTC 2247

QY 1321 CTCAGCCAAATGGATGCTGCTGCTTCTCCCTTCTTGGACCTCTAGCAGCTCAATAATG 1380

DB 2248 CTCAGCCAAATGGATGCTGCTGCTTCTCCCTTCTTGGACCTCTAGCAGCTCAATAATG 2307

QY 1381 TTACTCTCTTTGGACCTGATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGAAAT 1440

DB 2308 CTACTCTCTTTGGACCTGATCTTTAACTCTCTTGTAACTTTGTCTCTTCTCCAGAAATC 2367

QY 1441 GAAGCTGTAAAGCTACAGATGCTCTTCAAAATGGAACCCCA 1481

DB 2368 GAAGCTGTAAAACTACAAATGAGCCCAAGATGCAATCCAA 2408

RESULT 11
ADC38776

ID ADC38776 standard; cDNA; 2946 BP.

XX ADC38776;

XX 18-DEC-2003 (first entry)

XX Human cDNA encoding a secreted protein #63.

DE ss; gene; immune disorder; severe combined immunodeficiency; SCID;
XX autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;
KW lymphoid cell deficiency; osteoporosis; osteoarthritis;
KW peripheral nervous system disease; peripheral neuropathy;
KW Alzheimer's disease; Parkinson's disease; coagulation disorder;
KW inflammatory disease; systemic inflammatory response syndrome; SIRS;
KW ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;

KW hypersensitivity; regeneration; neural cell proliferation; fertility;
 KW tumour; chemokine; human; secreted protein.
 XX Homo sapiens.

OS US2002193567-A1.

PN 19-DEC-2002.

PD 02-APR-2002; 2002US-00114893.

XX 11-AUG-1995; 95US-005114014.

XX 05-APR-1996; 96US-00628364.

XX 19-APR-1996; 96US-00635311.

XX 07-JUN-1996; 96US-00659224.

XX 17-JUN-1996; 96US-00664596.

XX 09-JUL-1996; 96US-00677231.

XX 26-JUL-1996; 96US-00686878.

XX 23-AUG-1996; 96US-00701819.

XX 27-SEP-1996; 96US-00721488.

XX 27-SEP-1996; 96US-00721798.

XX 27-SEP-1996; 96US-00721923.

XX 27-SEP-1996; 96US-00721926.

XX 25-OCT-1996; 96US-00738367.

XX 30-OCT-1996; 96US-00739775.

XX 13-JAN-1997; 97US-00783395.

XX 10-APR-1997; 97US-00833823.

XX 02-JUN-1997; 97US-00867677.

XX 05-SEP-1997; 97US-00924838.

XX 06-OCT-1999; 99US-00413232.

PA (GEMY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;

PI Kelleher K;

XX WPI: 2003-657236/62.

DR P-PSDB; ADC38777.

XX Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and

PT BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,

PT useful for treating e.g. multiple sclerosis and rheumatoid arthritis.

XX Disclosure; SEQ ID NO 134; 412pp; English.

XX The invention relates to a protein comprising fully defined AZ302 1
 CC protein or BD127 1 6 protein. The polynucleotides are useful for
 CC expressing recombinant proteins for analysis and are also useful as
 CC chromosome markers or tags to identify chromosomes or to map related gene
 CC positions. The proteins are useful as amino acid supplement, carbon
 CC source, nitrogen source and carbohydrate source. The proteins are useful
 CC for treating various immune deficiencies and disorders (e.g. severe
 CC combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic
 CC reactions (e.g. asthma, myeloid or lymphoid cell deficiencies,
 CC osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.
 CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),
 CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory
 CC response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease),
 CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing
 CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or
 CC nerve growth or regeneration, for proliferating neural cells and for
 CC regenerating nerve and brain tissue, for inducing fertility and for
 CC inhibiting tumour growth. Proteins are also useful as chemokine for
 CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also
 CC useful as inhibitors of receptor/ligand interactions. The present
 CC sequence represents cDNA encoding a human secreted protein.

XX Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

XX Query Match 90.2%; Score 1335.4; DB 10; Length 2946;

XX Best Local Similarity 93.9%; Pred. No. 0;

		Matches 1390;	Conservative	0;	Mismatches	91;	Indels	0;	Gaps	0;
QY	1	ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTTACCCCTTTCGCTCACT	60							
DB	928	ATGGCCCTCCCTTATCATATTTTCTCTTTACTGTTTACCTCTTTCACTCTCACT	987							
QY	61	GCACCCCTCCATGCTGCTGTACAACAGTAGTCTCCCTTACCAAGAGTTCTATGAAGA	120							
DB	988	GCACCCCTCCATGCGCTGTATGACAGTAGTCTCCCTTACCAAGAGTTCTATGGAGA	1047							
QY	121	ACGCGCTTCTCTGGAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCTCC	180							
DB	1048	ATGACGCTCCCGAATATTTGATGCCCATCGTATAGAGTCTTCTAAGGGAATCTCC	1107							
QY	181	ACCTTCACCTGCCACACCCCATATGCCCCCACTGCTATAAATCTCTCCACTCTTTGCATG	240							
DB	1108	ACCTTCACCTGCCACACCCCATATGCCCCCACTGCTATAAATCTCTCCACTCTTTGCATG	1167							
QY	241	CATGCAATATCTATTTATTTGGACAGGGAATGATTAATCTTCTAGTTTCTCTGGAGACTT	300							
DB	1168	CATGCAATATCTATTTATTTGGACAGGGAATGATTAATCTTCTAGTTTCTCTGGAGACTT	1227							
QY	301	GGAGCCACTGCTGTTGGACTTACTTTCACCCATACCACTATGCTCTGATGGGGTGGAAAT	360							
DB	1228	GGAGTCACTGCTGTTGGACTTACTTTCACCCAACTGGTATGCTGATGGGGTGGAGTT	1287							
QY	361	CAAGGTCAAGCAAGAGAAAAAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGGACAT	420							
DB	1288	CAAGTCAAGCAAGAGAAAAAAGTAAAGGAAGTAAATCTCCCAACTGACCCCGGGTACAT	1347							
QY	421	AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCTCTCCGTACC	480							
DB	1348	GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCCGTACC	1407							
QY	481	CATACCTGCGCTGGTGGAGCTTATTTATACCACTCTCACTCGGCTCCATGAGGTCTCAGCC	540							
DB	1408	CATACCTGCGCTGGTGGAGCTTATTTATACCACTCTCACTCGGCTCCATGAGGTCTCAGCC	1467							
QY	541	CAAAACCTACTAACTGTTGGATGTGCTCCCCCTGCACCTTCAGGGGCAATACATTTCAATC	600							
DB	1468	CAAAACCTACTAACTGTTGGATGTGCTCCCCCTGCACCTTCAGGGGCAATATGTTTCAATC	1527							
QY	601	CTGTTCTCTGAAACAATGGAAACAATCTTCAGCACAGAAATAAACACCTCTCGTTTATGTA	660							
DB	1528	CCTGTACTCTGAAACAATGGAAACAATCTTCAGCACAGAAATAAACACCTCTCGTTTATGTA	1587							
QY	661	GGACCTCTGTTTCCCAATCTGGAAATAACCCATACCTCAAACTCACTCTGTGTAATAATTT	720							
DB	1588	GGACCTCTGTTTCCCAATCTGGAAATAACCCATACCTCAAACTCACTCTGTGTAATAATTT	1647							
QY	721	AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAACACCTCCACACGA	780							
DB	1648	AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAACCTCCACACCAA	1707							
QY	781	ATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCATTTGTTG	840							
DB	1708	ATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCTATCATTTGTTG	1767							
QY	841	AATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900							
DB	1768	AATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1827							
QY	901	ACTGAAACAAGATTTATACAATCATGCTACCTAAAGCCCAACAAACAAAGAGTACCCATT	960							
DB	1828	ACTGAAACAAGATTTATACAATCATGCTACCTAAAGCCCAACAAACAAAGAGTACCCATT	1887							
QY	961	CTTCCTTTCTTATCAGACAGGAGTCTAGGCAGACTAGGTACTGGCATTTGGCAGTATC	1020							
DB	1888	CTTCCTTTCTTATCAGACAGGAGTCTAGGCAGACTAGGTACTGGCATTTGGCAGTATC	1947							
QY	1021	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGTCATGACATGAAACAG	1080							
DB	1948	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGTCATGACATGAAACAG	2007							

QY 661 GGACCTCTGTTCCCAATCTGGAATAAACCACCTCAAACTCCTGCTGTGTAATAATTT 720
 Db 661 GGACCTCTGTTCCCAATCTGGAATAAACCACCTCAAACTCCTGCTGTGTAATAATTT 720
 QY 721 AGCAATCTATAGACACAAACAGCTCCCAATGATCAGGTGGTAAACACCTCCACACGA 780
 Db 721 AGCAATCTATACACAAACCAACTCCCAATGATCAGGTGGTAAACCTCCACACAA 780
 QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCATTTGTTG 840
 Db 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCATTTGTTG 840
 QY 841 AATGGCTCTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 Db 841 AATGGCTCTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 QY 901 ACTGAACAAGATTTATACATCTATGCTGTACCTAAAGCCCAACAAAGAGTACCCATT 960
 Db 901 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATT 960
 QY 961 CTTCCCTTTGTTATCAGAGCAGGAGTCTAGGACAGTACTAGTACTGGCATTGGCAGTATC 1020
 Db 961 CTTCCCTTTGTTATAGGAGCAGGAGTCTAGGACAGTACTAGTACTGGCATTGGCAGTATC 1020
 QY 1021 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGCTGACATGGACAG 1080
 Db 1021 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGCTGACATGGACAG 1080
 QY 1081 GTCACTGACTCCCTGGTCACTCTGCAAGATCAACTTAACCTCCCTAGCAGAGTAGTCTTT 1140
 Db 1081 GTGCCCACTCCCTGGTCACTCTGCAAGATCAACTTAACCTCCCTAGCAGAGTAGTCTTT 1140
 QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGAACTGTTTATTTT 1200
 Db 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTT 1200
 QY 1201 GGAGAAACGCTGTTTATTTATGTTAAATCAATCCAGAAATGCTCACTGAGAAAGTTAAAGAA 1260
 Db 1201 GGGGAAAGATGCTGTTTATTTATGTTAAATCAATCCGAAATGCTCACTGAGAAAGTTAAAGAA 1260
 QY 1261 ATTCGAGATCGAATACAAATGATAGACAGAGAGCTTCAAAACACCGCAACGCTGGGGCTC 1320
 Db 1261 ATTCGAGATCGAATACAAATGATAGACAGAGAGCTTCAAAACACCGCAACGCTGGGGCTC 1320
 QY 1321 CTCAGCCAATGGATGCGCTGGATTCTCCCTTCTTTAGGACCTCTAGCAGCTCTAATATTG 1380
 Db 1321 CTCAGCCAATGGATGCGCTGGATTCTCCCTTCTTTAGGACCTCTAGCAGCTCTAATATTG 1380
 QY 1381 TTACTCTCTTTGGACCTGATCTTTAACTCTCTGTTAAAGTTGTTGCTCTCTCCAGATT 1440
 Db 1381 CTACTCTCTTTGGACCTGATCTTTAACTCTCTGTTAAAGTTGTTGCTCTCTCCAGATT 1440
 QY 1441 GAAGCTGTAAAGCTACAGATGCTTACAAATGGAACCCCA 1481
 Db 1441 GAAGCTGTAAAGCTACAAATGGAACCCCAAGATGAGTCCAA 1481

RESULT 13
 ID ABN97948
 AC ABN97948;
 DT 01-AUG-2002 (first entry)
 DE Human retroviral HERV-7q env coding sequence.
 KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
 KW multiple sclerosis; ds.
 OS Human endogenous retrovirus.
 XX

PN WO967395-A1.
 XX 29-DEC-1999.
 XX 23-JUN-1999; 99WO-FR001513.
 XX 23-JUN-1998; 98FR-00007920.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA Alliel PM, Perin J, Rieger F;
 PI WPI; 2000-160587/14.
 XX
 XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
 PT for diagnosis, treatment and prevention of autoimmune and neurological
 PT diseases.
 XX
 XX Claim 3; Page 142-145; 225pp; French.
 XX
 XX The present invention relates to new nucleic acid sequences of human
 CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
 CC Regulatory elements associated with HERV-7q may alter expression of other
 CC genes (even remote genes) on the same chromosome, inducing immunological
 CC and/or neurological changes (which may be pathological or protective/
 CC curative). HERV-7q peptides can be used to improve efficiency of the
 CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
 CC sequences can be used in immunogenic or vaccinating compositions, for
 CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention
 XX
 SQ Sequence 2055 BP; 576 A; 574 C; 376 G; 529 T; 0 U; 0 Other;
 Query Match 90.1%; Score 1333.8; DB 3; Length 2055;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATCATCTATTTCTCTTTACTGTTCTCTTACCCCTTTCGCTCTCACT 60
 Db 391 ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTTACCCCTTTCGCTCTCACT 450
 QY 61 GCACCCCTCCATGCTGCTCAACCCAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
 Db 451 GCACCCCTCCATGCTGCTGATGACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 510
 QY 121 ACGGCGCTTCTCGAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCTC 180
 Db 511 ATGAGGGTCCCGAATATTTGATGCCCATCATATAGGAGTTTCTTCTAAGGGAATCTC 570
 QY 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATTAACCTGCGACCTTTTGCATG 240
 Db 571 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATTAACCTGCGACCTTTTGCATG 630
 QY 241 CATGCAATCTCATTTATTCGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 300
 Db 631 CATGCAATCTCATTTATTCGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 690
 QY 301 GGAGCCACTGCTGTTGGACTTATCTCACCATAACAGTATGCTGATGGGGTGGAAATT 360
 Db 691 GGAGTCACTGCTGTTGGACTTATCTCACCATAACAGTATGCTGATGGGGTGGAAATT 750
 QY 361 CAAGGTGAGGACAGAGAAAAACAAGTAAAGGACCACTCTCCCACTGACCCGGGACAT 420
 Db 751 CAAGATCAGGACAGAGAAAAACAAGTAAAGGACCACTCTCCCACTGACCCGGGACAT 810
 QY 421 AGCACCTCTAGCCCTTCAAGAGGACTAGTTCTCTCAAAATCTACATGAACCCCTCGTACC 480
 Db 811 GGCACCTCTAGCCCTTCAAGAGGACTAGATCTCTCAAAATCTACATGAACCCCTCGTACC 870
 QY 481 CATACTCGCTGCTGAGCTATTTTAATACCAACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540

Db 871 CATACTCGCTGGTAGCCTATTTATACCACTCTCACTGGGCTCCATGAGGTCTCGGC 930
Qy 541 CAAAACCCCTACTAACTGTTGGATGCGCTCCCTGCACTTCAGGCCATACATTTCAATC 600
Db 931 CAAAACCCCTACTAACTGTTGGATGCGCTCCCTGCACTTCAGGCCATATGTTTCAATC 990
Qy 601 CTGTTCTCTGAACATGGAACAACTTCAGCAGAGAAATAAACACCACTTCCTGTTTAGTA 660
Db 991 CTGTACTCTGAAACAACTGGAACAACTTCAGCAGAGAAATAAACACCACTTCCTGTTTAGTA 1050
Qy 661 GGACCTCTGTTTCCAACTCTGAAATAACCCATACCTCAAACTCTCACCTGTGTAAATTT 720
Db 1051 GGACCTCTGTTTCCAACTCTGAAATAACCCATACCTCAAACTCTCACCTGTGTAAATTT 1110
Qy 721 AGCAATATATAGACAAACACAGCTCCCAATGATCAGGTGGGTAAACACCTCCCAACGA 780
Db 1111 AGCAATATATAGACAAACACCACTCCCAATGATCAGGTGGGTAAACCTCCCAACGA 1170
Qy 781 ATAGTCTGCTACCTCAGGAATATTTTGTGTTGCTGCTGAGCTCAGCCTATCATTTGTTG 840
Db 1171 ATAGTCTGCTACCTCAGGAATATTTTGTGTTGCTGCTGAGCTCAGCCTATCATTTGTTG 1230
Qy 841 AATGCTCTTTCAGAACTATGCTCTCTCTCATTTAGTGCCTCTATGACCATCTAC 900
Db 1231 AATGCTCTTTCAGAACTATGCTCTCTCTCATTTAGTGCCTCTATGACCATCTAC 1290
Qy 901 ACTGAACAAGATTTATACAAATGCTGTAACCTAAGCCCCCAACAAAGAGTACCCATT 960
Db 1291 ACTGAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCCAACAAAGAGTACCCATT 1350
Qy 961 CTTCTCTTTGTTATCAGACGAGGCTAGGAGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGG 1020
Db 1351 CTTCTCTTTGTTATAGGAGGAGGCTAGGAGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGG 1410
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 1411 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1470
Qy 1081 GTCACGTACTCCTGTGCTACCTTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCCTT 1140
Db 1471 GTGCGGACTCCTGCTGCTACCTTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCCTT 1530
Qy 1141 CAAATCGAGAGCTTTAGACTGCTTAACCGCAAGAGGGGACCTGTTATTTTA 1200
Db 1531 CAAATCGAGAGCTTTAGACTGCTTAACCGCTGAAGAGGGGACCTGTTATTTTA 1590
Qy 1201 GGAGAAAGACGCTGTTATTTATGTTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA 1260
Db 1591 GGGAAGAAATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1650
Qy 1261 ATTGAGATGGAATACAAATGATAGAGAGGAGCTTCAAAACACCGAAGCTGGGCGCTC 1320
Db 1651 ATTGAGATGGAATACAAATGATAGAGAGGAGCTTCAAAACACCTGGACCTGGGCGCTC 1710
Qy 1321 CTGACCAATGATGCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATG 1380
Db 1711 CTGACCAATGATGCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATG 1770
Qy 1381 TTACTCTCTTTGGACCTCTGATCTTTTAACTCTCTTTTAAAGTTTGTCTCTTCCAGAAAT 1440
Db 1771 CTACTCTCTTTGGACCTCTGATCTTTTAACTCTCTTTTAACTTTGTCTCTTCCAGAAAT 1830
Qy 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 1831 GAAGCTGTAAAGCTACAAATGGAACCCCAAGATGAGTCCAA 1871

RESULT 14
ABN97927
ID ABN97927 standard; DNA; 2599 BP.
XX
AC ABN97927;

XX 01-AUG-2002 (first entry)
DT Human retroviral sequence HERV 7 env.
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis; ds.
XX Human endogenous retrovirus.
XX WO9967395-A1.
XX 29-DEC-1999.
XX 23-JUN-1999; 99WO-FR001513.
XX 23-JUN-1998; 98FR-00007920.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Alliel PM, Perin J, Rieger F;
XX WPI; 2000-160587/14.
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
PT for diagnosis, treatment and prevention of autoimmune and neurological
PT diseases.
XX Claim 1; Page 128-129; 225pp; French.
PS The present invention relates to new nucleic acid sequences of human
CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
CC Regulatory elements associated with HERV-7q may alter expression of other
CC genes (even remote genes) on the same chromosome, inducing immunological
CC and/or neurological changes (which may be pathological or protective/
CC curative). HERV-7q peptides can be used to improve efficiency of the
CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention
XX Sequence 2599 BP; 744 A; 718 C; 495 G; 642 T; 0 U; 0 Other;
SQ Query Match 90.1%; Score 1333.8; DB 3; Length 2599;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
Qy 1 ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTCTCTTACCCCTTTCTCACT 60
Db 929 ATGGCCCTCCCTTATCATATTTTCTTTACTGTCTTTTACCCTCTTCACTCTCACT 988
Qy 61 GCACCCCTCCATGCTGTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 989 GCACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCAAGAGTTTCTATGGAGA 1048
Qy 121 AGCGGCTCTCGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 1049 ATGACGCTGCCGAAATATTGATGCCCATCATATAGGAGTTTCTAAGGGAACCTCC 1108
Qy 181 ACCTTCACTGCCACACCATATGCTCCGCAACTGCTAACTGCACTCTTTGCAATG 240
Db 1109 ACCTTCACTGCCACACCATATGCTCCGCAACTGCTATCACTCTGCACTCTTTGCAATG 1168
Qy 241 CATGCAAAATCACTATTATGGACAGGGAATGATTAATCTTAGTTGCTCTGAGGACTT 300
Db 1169 CATGCAAAATCACTATTATGGACAGGGAATGATTAATCTTAGTTGCTCTGAGGACTT 1228
Qy 301 GGAGCCTGCTGTGTTGAGCTTACTTCCACCATACAGTATCTCTGATGGGTGGAATT 360
Db 1229 GGAGCCTGCTGTGTTGAGCTTACTTCCACCATACAGTATCTCTGATGGGTGGAATT 1288

QY 361 CAAGGTCAGCAGAGAAAAACAGTAAGGAGCAATCTCCCAACTGACCCGGGACAT 420
Db 1289 CAAGATCAGCAGAGAAAAACATGTAAGAAGTAATCTCCCACTCACCAGGATAT 1348
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
Db 1349 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 1408
QY 481 CATACTCGCTGTGAGCCATTAATATACACCCCTCACTCGGCTCATGAGGTCTCAGCC 540
Db 1409 CATACTCGCTGTGAGCCATTAATATACACCCCTCACTCGGCTCATGAGGTCTCAGCC 1468
QY 541 CAAAACCCCTACTACTGTTGGATGCTCCCTCCCTGCACCTTCAAGGCATATTTCAATC 600
Db 1469 CAAAACCCCTACTACTGTTGGATGCTCCCTCCCTGCACCTTCAAGGCATATTTCAATC 1528
QY 601 CTGTGTTCTGAAACAATGGAACAATCTCAGCACAGAAATAAACACCACTTCCGTTTAGTA 660
Db 1529 CTGTGTTCTGAAACAATGGAACAATCTCAGCACAGAAATAAACACCACTTCCGTTTAGTA 1588
QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCAGCTGTGTAAATTT 720
Db 1589 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCAGCTGTGTAAATTT 1648
QY 721 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
Db 1649 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 1708
QY 781 ATAGTCGCTACCCCTCAGGAATATTTTGTGTTGTTACCTCAGCTATCATTTGTTG 840
Db 1709 ATAGTCGCTACCCCTCAGGAATATTTTGTGTTGTTACCTCAGCTATCATTTGTTG 1768
QY 841 AATGGCTCTTCAAGATCTATGTTCTCTCATTTAGTGGCCCTATGACCATCTAC 900
Db 1769 AATGGCTCTTCAAGATCTATGTTCTCTCATTTAGTGGCCCTATGACCATCTAC 1828
QY 901 ACTGAAACAGTTTATACATATCTGTACCTTAAGCCCAACAAAGAGTACCCATT 960
Db 1829 ACTGAAACAGTTTATACATATCTGTACCTTAAGCCCAACAAAGAGTACCCATT 1888
QY 961 CTTCCTTTTGTATCAGACGAGGAGTCTAGGCAGACTAGTACTGCGATTGGCAGTATC 1020
Db 1889 CTTCCTTTTGTATCAGACGAGGAGTCTAGGCAGACTAGTACTGCGATTGGCAGTATC 1948
QY 1021 ACAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGTGACATGGAACAG 1080
Db 1949 ACAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGTGGAACATGGAACGG 2008
QY 1081 GTCACTGACTCCCTGGTCACCTTCAAGATCACTTAACCTCCCTAGCAGAGTAGTCTCT 1140
Db 2009 GTCCCGACTCCCTGGTCACCTTCAAGATCACTTAACCTCCCTAGCAGAGTAGTCTCT 2068
QY 1141 CAAAATCGAAGAGCTTTAGACTCTGCTAACCGCCAAAAGAGGGGAACTGTTTATTTTA 1200
Db 2069 CAAAATCGAAGAGCTTTAGACTCTGCTAACCGCTGAAGAGGGGAACTGTTTATTTTA 2128
QY 1201 GGAGAAGAACCTGTTATTTATTAATCAATPCCAGATTTGCTAGAGAAAGTTAAAGAA 1260
Db 2129 GGGAAGAAATGCTGTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 2188
QY 1261 ATTCGAGATCGAATACATGATGAGCAGAGAGGCTTCAAAACCGGACCGCTGGGGCTC 1320
Db 2189 ATTCGAGATCGAATACACGATGAGCAGAGAGGCTTCAAAACCGGACCGCTGGGGCTC 2248
QY 1321 CTCAGCCAAATGATGCTGGTCTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTTG 1380
Db 2249 CTCAGCCAAATGATGCTGGTCTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTTG 2308
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCCCTGTTGTTAGTTGTTCTTCCAGAAAT 1440
Db 2309 CTACTCTCTTTGGACCCCTGTATCTTTAACTCCCTGTTGTTAGTTGTTCTTCCAGAAAT 2368

QY 1441 GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAAACCCCA 1481
Db 2369 GAAGCTGTAAAGCTACAAATGGAAAGCCCAAGATGCAATGCCAA 2409
RESULT 15
ADS30988
ID ADS30988 standard; DNA; 8523 BP.
XX AC ADS30988;
XX DT 18-NOV-2004 (first entry)
XX Human genome high complexity repeat found in the HIRA gene #21.
XX Human; ds;
KW histone cell cycle regulation defective, S. cerevisiae homologue A; HIRA;
KW high complexity repeat; in situ hybridisation; Southern blot;
KW chromosome breakpoint; inherited genetic disease; neoplastic disorder;
KW chromosome 22; DiGeorge syndrome; Velo-Cardio-facial syndrome.
XX Homo sapiens.
OS US2003224356-A1.
XX PN 04-DEC-2003.
XX PD 14-MAY-2001; 2001US-00854867.
XX PF 16-MAY-2000; 2000US-00573080.
XX PR (KNOL/) KNOLL J H M.
XX PA (ROGA/) ROGAN P K.
XX PA Knoll JHM, Rogan PK;
PI WPI; 2002-062378/08.
DR
XX Single copy genomic hybridization probes for detecting specific nucleic
PT acid sequences in sample by in situ hybridization useful for detection of
PT acquired or inherited genetic diseases.
XX Example 1; SEQ ID NO 21; 30pp; English.
XX The invention relates to a nucleic acid hybridisation probe comprising a
CC labelled, single copy nucleic acids of at least 50 nucleotides, which
CC will hybridise to a deduced single copy sequence interval in target
CC nucleic acid (TNA) of known sequence. The single copy sequence is deduced
CC by comparing the target nucleic acid (e.g. a disease causing gene) with a
CC collection of high and low complexity repeat sequences as found in the
CC genome of the organism from containing the target nucleic acid. The probe
CC is generated by PCR on the target sequence. The probe is essentially free
CC of blocking nucleic acid sequences which will hybridise to repeat
CC sequences within the genome of which the TNA is a part, and is labelled
CC with a label selected from fluorochrome-responsive labels, fluorochromes,
CC calorimetric chemical, conjugated proteins, antibodies, antigens and
CC their mixtures. The probe is useful in a hybridisation method, where the
CC hybridisation method is from in situ hybridisation, Southern blot, and
CC other methods in which nucleic acid is immobilised, where the method
CC further comprises selecting a single copy nucleic acid which will
CC hybridise to a duplcon or triplicon sequence domain. The probe is useful
CC for determining the existence of previously unknown repeat sequence
CC families in a genome. The method comprises reacting a labelled probe with
CC the genome, causing the probe to hybridise and ascertaining if the probe
CC hybridises to the genome at more than three preferably ten different
CC locations as a determination of new repeat sequence family, where the
CC determining step comprises selecting the single copy sequence from a
CC duplcon or triplicon sequence domain. The probe is useful for
CC determining a chromosome breakpoint and is useful in the fields for
CC cytogenetics and molecular genetics for determining the presence of
CC specific nucleic acid sequences in a sample of eukaryotic origin, e.g.
CC the probes may be used to analyse specific chromosomal locations by in
CC situ hybridisation as a detection of acquired or inherited genetic

CC diseases especially for detection of genetic or neoplastic disorders.
CC Unlike prior art techniques, the probe permits more precise chromosomal
CC breakpoint determinations by in situ hybridisation. The genomic sequence
CC comprising the human HIRA gene (histone cell cycle regulation defective,
CC S. cerevisiae, homologue A) was analysed for single copy sequence
CC intervals for use as probes of the invention. HIRA is located on
CC chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and
CC Velo-Cardio-facial syndromes. The present sequence is a high complexity
CC repeat found within the human genome used to analyse the HIRA gene for
CC repeat regions. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030224356.
XX

Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;

Query Match	90.1%;	Score 1333.8;	DB 7;	Length 8523;
Best Local Similarity	93.8%;	Pred. No. 0;		
Matches 1389;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;
QY	1	ATGGCCCTCCCTTATCATACTTTTCTTTTACGTGTTCTTACCCGCTTTTCGCTCTCACT	60	
DB	6849	ATGGCCCTCCCTTATCATATTTTCTTTTACTGTTCTTTTACCCCTCTTTTCACTCTCACT	6908	
QY	61	GCACCCCTCCATGCTGCTGTACACACGAGTCCCTTACCAAGAGTTTCTATGAAGA	120	
DB	6909	GCACCCCTCCATGCGCGCTGTATGACACGAGTAGTCCCTTACCAAGAGTTTCTATGGAGA	6968	
QY	121	ACGCGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC	180	
DB	6969	ATGCAGGTTCCCGAATAATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC	7028	
QY	181	ACCTTCACTGCCACACCCATATGCCCCGAACTGCTATAAATCTGCGCACTTTTGCAATG	240	
DB	7029	ACCTTCACTGCCACACCCATATGCCCCGAACTGCTATAAATCTGCGCACTTTTGCAATG	7088	
QY	241	CATGCAAACTCATTTATGGACAGGAAATGATTAATCTAGTTGTCTGAGGACATT	300	
DB	7089	CATGCAAACTCATTTATGGACAGGAAATGATTAATCTAGTTGTCTGAGGACATT	7148	
QY	301	GGAGCCACTCTCTGTTGGACTTACTTACCACCAACTGGTATGTCGTATGGGGTGGAGTT	360	
DB	7149	GGAGTCACTCTCTGTTGGACTTACTTACCACCAACTGGTATGTCGTATGGGGTGGAGTT	7208	
QY	361	CAAGGTCAGGCAAGAGAAAAAAGTAAGGAGCAATCTCCCACTGACCCGGGACAT	420	
DB	7209	CAAGATCAGGCAAGAGAAAAAAGTAAGGAGCAATCTCCCACTGACCCGGGACAT	7268	
QY	421	AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAATACATGAAACCCCTCCGTACC	480	
DB	7269	GGCACCTCTAGCCCTACAAAGGACTAGTTCTCTCAAAATACATGAAACCCCTCCGTACC	7328	
QY	481	CATACTCGCTGTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGCTCAGCC	540	
DB	7329	CATACTCGCTGTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGCTCAGCC	7388	
QY	541	CAAAACCCCTACTAATCTGATGAGTGTCTCCCTGCACTTCAGGCCATACATTCCAATC	600	
DB	7389	CAAAACCCCTACTAATCTGATGAGTGTCTCCCTGCACTTCAGGCCATATGTTTCAATC	7448	
QY	601	CCTGTTCTGAAACAACTTCCAGACAGAAATAAACACCACTTCCGTTTTAGTA	660	
DB	7449	CCTGTTCTGAAACAACTTCCAGACAGAAATAAACACCACTTCCGTTTTAGTA	7508	
QY	661	GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACTCAGCTGTGTAATAATT	720	
DB	7509	GGACCTCTTGTTTCCAATCTGGAAATAACCCATACCTCAAACTCAGCTGTGTAATAATT	7568	
QY	721	AGCAATACTATAGACACACACGCTCCCAATGCAATCAGGTGGGTAAACCTCCACACGA	780	
DB	7569	AGCAATACTATACACACCACTCCCAATGCAATCAGGTGGGTAACTCTCCACACAA	7628	
QY	781	ATAGTCTGCTCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG	840	

Db	7629	ATAGTCTGCTCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCCTATCGTTGTTG	7688	
QY	841	AATGGCTCTTCAGAACTATGCTTCTCTCTCAATTTCTAGTGCCCTATGACCATCTAC	900	
Db	7689	AATGGCTCTTCAGAACTATGCTTCTCTCTCAATTTCTAGTGCCCTATGACCATCTAC	7748	
QY	901	ACTGAAACAAGATTATACAATCATGTCGTACCTAAAGCCCAACAAAAAGAGTACCCATT	960	
Db	7749	ACTGAAACAAGATTATACAAGTATGTCATATCTAAAGCCCAACAAAAAGAGTACCCATT	7808	
QY	961	CTTCCTTTTGTATCAGAGCAGAGTCTAGGACAGACTAGGTACTGCGATGGCAGTATC	1020	
Db	7809	CTTCCTTTTGTATAGGAGCAGAGTCTAGGTACTGCGATGGCAGTATGGCGGTATC	7868	
QY	1021	ACAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAAATGTGACATGGAACAG	1080	
Db	7869	ACAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAAATGTGGAACATGGAACGG	7928	
QY	1081	GTCACTGACTCCCTGGTCACTTGCAGAGATCAACTTAACCTCCCTAGCAGAGTAGTCCTT	1140	
Db	7929	GTGCGGACTCCCTGGTCACTTGCAGAGATCAACTTAACCTCCCTAGCAGAGTAGTCCTT	7988	
QY	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTA	1200	
Db	7989	CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTA	8048	
QY	1201	GGAGAGAACCTGTTTATTTATTTAAATCCAGAAATTTGTCACTGAGAAAGTTAAAGAA	1260	
Db	8049	GGGGAAGAAATGCTGTTTATTTATTTAAATCCAGAAATTTGTCACTGAGAAAGTTAAAGAA	8108	
QY	1261	ATTGAGATCGAATACAAATGTAGACAGAGAGCTTCAAAAACACCGAACGCTGGGCGCTC	1320	
Db	8109	ATTGAGATCGAATACAAATGTAGACAGAGAGCTTCAAAAACACCGAACGCTGGGCGCTC	8168	
QY	1321	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380	
Db	8169	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	8228	
QY	1381	TTACTCTCTTTGGACCTCTATCTTTAACTCTCTTTTAAAGTTTGTCTCTTCCAGAAAT	1440	
Db	8229	CTACTCTCTTTGGACCTCTATCTTTAACTCTCTTTTAACTCTCTTGTAACTTTGTCTCTTCCAGAAATC	8288	
QY	1441	GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA	1481	
Db	8289	GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAAGTCCAA	8329	
RESULT 16				
ABN97929				
ID	ABN97929	standard; DNA; 10499 BP.		
XX	AC	ABN97929;		
XX	XX			
DT	01-AUG-2002	(first entry)		
XX	DE	Human retroviral sequence HERV-7q.		
XX	XX	Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;		
KW	XX	multiple sclerosis; ds.		
OS	XX	Human endogenous retrovirus.		
XX	XX	WO9967395-A1.		
PN	XX	29-DEC-1999.		
PD	XX	23-JUN-1999; 99WO-FR001513.		
PF	XX	23-JUN-1998; 98FR-00007920.		
XX	XX	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.		
PA	XX	Alliel PM, Perin J, Rieger F;		
PI	XX			

XX WPI; 2000-160587/14.

XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used

PT for diagnosis, treatment and prevention of autoimmune and neurological

PT diseases.

XX Claim 3; Fig 1; 225pp; French.

XX The present invention relates to new nucleic acid sequences of human

CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.

CC Regulatory elements associated with HERV-7q may alter expression of other

CC genes (even remote genes) on the same chromosome, inducing immunological

CC and/or neurological changes (which may be pathological or protective/

CC curative). HERV-7q peptides can be used to improve efficiency of the

CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding

CC sequences can be used in immunogenic or vaccinating compositions, for

CC protection against autoimmune diseases, particularly multiple sclerosis.

CC The peptides may also be used (by sequence comparison) to detect/identify

CC endogenous retroviruses that are abnormally expressed in cancer,

CC neuropathologies or other autoimmune diseases. The present sequence was

CC used to illustrate the invention

XX

SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 U; 0 Other;

Query Match 90.1%; Score 1333.8; DB 3; Length 10499;

Best Local Similarity 93.8%; Pred. No. 0;

Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATCTTTCTCTTACTGTTCTTTACCCCTTCGCTCTCACT 60

DB 7879 ATGGCCCTCCCTTATCATATTTTCTTACTGTTCTTTTACCTCTTTCACCTCACT 7938

QY 61 GCACCCCTCCATGCTGCTGTACAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 120

DB 7939 GCACCCCTCCATGCTGCTGTATGACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 7998

QY 121 ACGGGCTCTTGGAAATATGATGCCCATCATATATAGGAGTTTATCTAAGGGAACCTCC 180

DB 7999 ATGCAGCGTCCCGAAATATATGATGCCCATCATATAGGAGTTTCTTCTAAGGGAACCTCC 8058

QY 181 ACCTTCACTGCCACACCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTTGCATG 240

DB 8059 ACCTTCACTGCCACACCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTTGCATG 8118

QY 241 CATGCAAACTACTCATATATGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 300

DB 8119 CATGCAAACTACTCATATATGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 8178

QY 301 GGAGCCACTGCTGTTGGACTTACTTCAACCATACCATGCTGCTGATGGGGTGGAAAT 360

DB 8179 GGAGTCACTGCTGTTGGACTTACTTCAACCAACTGCTGATGCTGATGGGGTGGAGTT 8238

QY 361 CAAGGTCAAGCAAGAGAAACAAAGTAAAGGAACAATCTCCCAACTGACCCGGGACAT 420

DB 8239 CAAGATCAGGCAAGAGAAACAAATGTAAGAGAGTATCTCCCAACTCACCCTGGGTACAT 8298

QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480

DB 8299 GGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 8358

QY 481 CATACTGCTGGTGGAGCTATTTAATACCACTACCATGCTGGCTCCATGAGTCTCAGCC 540

DB 8359 CATACTGCTGGTGGAGCTATTTAATACCACTACCATGCTGGCTCCATGAGTCTCAGCC 8418

QY 541 CAAACCCCTACTACTGTTGGATGTCCTCCCTCGCACTTCAGGCCATACATTTCAATC 600

DB 8419 CAAACCCCTACTACTGTTGGATGTCCTCCCTCGCACTTCAGGCCATATGTTTCAATC 8478

QY 601 CCTGTTCTTGAACTAGGAACTTTCAGCAGAGAAATAACAACACTTCCTGTTTGTAGTA 660

DB 8479 CCTGTACCTGAACAACTTTCAGCAGAGAAATAACAACACTTCCTGTTTGTAGTA 8538

QY 661 GGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAAACTCACCCTGTGTAAATTT 720

DB 8539 GGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAAACTCACCCTGTGTAAATTT 8598

QY 721 AGCAATACCTATAGACACACACACAGCTCCCAATCATCATCAGCTGGTAAACACTCCACACAGA 780

DB 8599 AGCAATACCTATAGACACACACACAGCTCCCAATCATCATCAGCTGGTAAACACTCCACACAA 8658

QY 781 ATAGTCTGCTACCTCCCTCAGGAATAATTTTGTGTGTGCTACCTCAGCCCTATCATTTGTTG 840

DB 8659 ATAGTCTGCTACCTCCCTCAGGAATAATTTTGTGTGTGCTACCTCAGCCCTATCATTTGTTG 8718

QY 841 AATGGCTCTTCAGAACTATATGCTTCTCTCATCTTATGTCCTTCTAGTCCCTTATGACCATCTAC 900

DB 8719 AATGGCTCTTCAGAACTATATGCTTCTCTCATCTTATGTCCTTCTAGTCCCTTATGACCATCTAC 8778

QY 901 ACTGAACAGATTTATACAACTATCATGCTGCTACCTTAAGCCCCACAAACAAAGAGTACCCATT 960

DB 8779 ACTGAACAGATTTATACAACTATCATGCTGCTACCTTAAGCCCCACAAACAAAGAGTACCCATT 8838

QY 961 CTTCCCTTTTGTATCAGACGAGGAGTCTAGGAGAGTCTAGGTACTAGGTACTGGCATTTGGCAGTATC 1020

DB 8839 CTTCCCTTTTGTATCAGACGAGGAGTCTAGGAGAGTCTAGGTACTAGGTACTGGCATTTGGCAGTATC 8898

QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAACAG 1080

DB 8899 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAACAG 8958

QY 1081 GTCACTGACTCCCTGCTGCTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTT 1140

DB 8959 GTCCCGACTCCCTGCTGCTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTT 9018

QY 1141 CAAATCGAAGAGCTTTAGACTTTGCTAAACCGCAAGAGGGGGAACCTGTTTATTTTAA 1200

DB 9019 CAAATCGAAGAGCTTTAGACTTTGCTAAACCGCAAGAGGGGGAACCTGTTTATTTTAA 9078

QY 1201 GGAGAAAGACGCTGTTTATGTTTATCAATCAGAAATTTGCTACTGAGAAAGTTAAAGAA 1260

DB 9079 GGAGAAAGACGCTGTTTATGTTTATCAATCAGAAATTTGCTACTGAGAAAGTTAAAGAA 9138

QY 1261 ATTGAGATCGAATCAATGATAGACAGAGGAGCTTCAAAACACCGCAACCTGGGGCTC 1320

DB 9139 ATTGAGATCGAATCAATGATAGACAGAGGAGCTTCAAAACACCGCAACCTGGGGCTC 9198

QY 1321 CTCAGCCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

DB 9199 CTCAGCCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 9258

QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGATT 1440

DB 9259 CTACTCTCTTTGGACCCCTGTATCTTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGATT 9318

QY 1441 GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAAACCCCA 1481

DB 9319 GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAAACCCCA 9359

RESULT 17

ABL61744

ID ABL61744 standard; DNA; 56093 BP.

XX ABL61744;

AC ABL61744;

XX

DT 15-MAY-2002 (first entry)

XX

DE Colon adenocarcinoma related gene sequence SEQ ID NO:81.

XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX gene; ds.

OS Homo sapiens.

XX WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
XX 18-SEP-2000; 2000US-0233133P.
XX 18-SEP-2000; 2000US-0233161P.
XX 20-SEP-2000; 2000US-0234009P.
XX 20-SEP-2000; 2000US-0234034P.
XX 20-SEP-2000; 2000US-0234052P.
XX 22-SEP-2000; 2000US-0234509P.
XX 22-SEP-2000; 2000US-0234567P.
XX 25-SEP-2000; 2000US-0234923P.
XX 25-SEP-2000; 2000US-0234924P.
XX 25-SEP-2000; 2000US-0235077P.
XX 25-SEP-2000; 2000US-0235082P.
XX 25-SEP-2000; 2000US-0235134P.
XX 25-SEP-2000; 2000US-0235280P.
XX 26-SEP-2000; 2000US-0235637P.
XX 26-SEP-2000; 2000US-0235638P.
XX 27-SEP-2000; 2000US-0235711P.
XX 27-SEP-2000; 2000US-0235720P.
XX 27-SEP-2000; 2000US-0235840P.
XX 27-SEP-2000; 2000US-0235863P.
XX 28-SEP-2000; 2000US-0236028P.
XX 28-SEP-2000; 2000US-0236032P.
XX 28-SEP-2000; 2000US-0236033P.
XX 28-SEP-2000; 2000US-0236034P.
XX 28-SEP-2000; 2000US-0236109P.
XX 28-SEP-2000; 2000US-0236111P.
XX 29-SEP-2000; 2000US-0236842P.
XX 29-SEP-2000; 2000US-0236891P.
XX 02-OCT-2000; 2000US-0237172P.
XX 02-OCT-2000; 2000US-0237173P.
XX 02-OCT-2000; 2000US-0237278P.
XX 02-OCT-2000; 2000US-0237294P.
XX 02-OCT-2000; 2000US-0237295P.
XX 02-OCT-2000; 2000US-0237316P.
XX 03-OCT-2000; 2000US-0237425P.
XX 03-OCT-2000; 2000US-0237598P.
XX 03-OCT-2000; 2000US-0237604P.
XX 03-OCT-2000; 2000US-0237606P.
XX 03-OCT-2000; 2000US-0237608P.
XX 01-NOV-2000; 2000US-0244867P.
XX 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 81; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result

CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms's
CC tumour
XX
XX Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;
SQ
Query Match 90.1%; Score 1333.8; DB 6; Length 56093;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCCCTTATCATACTTTTCTTACTGCTTCTTACTGCTTCTTACCCCTTCGCTCTCACT 60
DB 35879 ATGCGCCCTCCCTTATCATATTTTCTTACTGCTTCTTACTGCTTCTTACCCCTTCGCTCTCACT 35938
QY 61 GCACCCCTCCATGCTGCTGTACCAACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 35939 GCACCCCTCCATGCGGCTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 35998
QY 121 ACGCGGCTTCCTGAAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB 35999 ATGCAGCGTCCCGGAAATATTTGATGCCCATCATATAGGAGTTTCTTAAAGGGAACCTCC 36058
QY 181 ACCTTCAGTCCGCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 240
DB 36059 ACCTTCAGTCCGCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 36118
QY 241 CATGCAATACTCATTTATTGGACAGGGAATAATGATTAATCTTAGTTGCTTGGAGGACTT 300
DB 36119 CATGCAATACTCATTTATTGGACAGGGAATAATGATTAATCTTAGTTGCTTGGAGGACTT 36178
QY 301 GGAGCACTGCTGTGGAGCTTACTTACCATATACCAGTATGCTGATGGGGTGGAAATT 360
DB 36179 GGAGTCACTGCTGTGGAGCTTACTTACCCCAATGCTATGCTGATGGGGTGGAGTT 36238
QY 361 CAAGGTCAAGCAAGAGAAACAAAGTAAGGAAGCAATCTCCCAACTCACCCGGGAGCAT 420
DB 36239 CAAGATCAGGCAAGAGAAACAAAGTAAGGAAGTAACTCTCCCAACTCACCCGGGTACAT 36298
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTATACGAAACCTCCGTACC 480
DB 36299 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTATACGAAACCTCCGTACC 36358
QY 481 CATACTCGCTGCTGAGGCTTAAATACCAACCTCACTCCGCTCCATGAGGTCTCAGCC 540
DB 36359 CATACTCGCTGCTGAGGCTTAAATACCAACCTCACTCCGCTCCATGAGGTCTCAGCC 36418
QY 541 CAAAACCTTACTAAGTGTGGATGTCCTCCCTCCCTGCACTTCAGGCCATACATTTCAATC 600
DB 36419 CAAAACCTTACTAAGTGTGGATGTCCTCCCTCCCTGCACTTCAGGCCATATGTTTCAATC 36478
QY 601 CCTGTTCTGAAACAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
DB 36479 CCTGTACCTGAAACAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 36538
QY 661 GGACCTCTGTTTCAATCTGGAATAACCACTACCTCAAACTCACCTGTGTAAATTT 720
DB 36539 GGACCTCTGTTTCAATCTGGAATAACCACTACCTCAAACTCACCTGTGTAAATTT 36598
QY 721 AGCAATCTATAGACAAACAGCTCCCAATCATCATCAGGTGGGTAAACACCTCCCAACCA 780
DB 36599 AGCAATCTATAGACAAACAGCTCCCAATCATCATCAGGTGGGTAAACCTCCCAACCA 36658
QY 781 ATAGTCTGCTACCTCAGGAATAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 36659 ATAGTCTGCTACCTCAGGAATAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 36718
QY 841 AATGCTCTTCAAGATCTATGCTTCTCTCTATCTTCTAGTCCCTCATGACCATCTAC 900
DB 36719 AATGCTCTTCAAGATCTATGCTTCTCTCTCTATCTTCTAGTCCCTCATGACCATCTAC 36778

QY 901 ACTGAACAAGATTTATACAAATCATGTCGTACCTAAGCCCAACCAACAAAGAGTACCCATT 960
 DB 36779 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATT 36838
 QY 961 CTTCCCTTTTGTATACAGAGCAGAGTCTAGGAGACTAGGTACTGGCATTGGAGTATC 1020
 DB 36839 CTTCCCTTTTGTATAGGAGCAGAGTCTAGGTGCACTAGGTACTGGCATTGGCGGTATC 36898
 QY 1021 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATTAATGTGACATGGACAG 1080
 DB 36899 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATTAATGTGAGCATGGACGG 36958
 QY 1081 GTCACTGACCTCCCTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGAGTATGCTTT 1140
 DB 36959 GTGCCGACCTCCCTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGAGTATGCTTT 37018
 QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGAACTGTTTATTTTAA 1200
 DB 37019 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAGAGGGGAACTGTTTATTTTAA 37078
 QY 1201 GGAGAAAGAGCTTTATTTATTTATTAATCAATCCAGAAATTCCTGAGAAAGTTAAAGAA 1260
 DB 37079 GGGGAAGAATGCTGTTTATTTATTTATTAATCAATCCAGAAATTCCTGAGAAAGTTAAAGAA 37138
 QY 1261 ATTGAGATCGAATACAAATGATGAGAGAGAGGCTTCAAAACACCGCAACGCTGGGGCTC 1320
 DB 37139 ATTGAGATCGAATACAAATGATGAGAGAGAGGCTTCAAAACACCGCAACGCTGGGGCTC 37198
 QY 1321 CTCAGCAATGAGTCCCTGGTCTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATG 1380
 DB 37199 CTCAGCAATGAGTCCCTGGTCTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATG 37258
 QY 1381 TTACTCTCTTTGAGACCTGATCTTTAACTCTCTGTTTAAAGTTTGTCTCTCCAGAAAT 1440
 DB 37259 CTACTCTCTTTGAGACCTGATCTTTAACTCTCTGTTTAAAGTTTGTCTCTCTCCAGAAAT 37318
 QY 1441 GAGCTCTAAAGCTACAGAGTGTCTTCAAAATGGAACCCCA 1481
 DB 37319 GAGCTCTAAAGCTACAGAGTGTCTTCAAAATGGAACCCCA 37359

RESULT 18
 ID AAF55630 standard; DNA; 2781 BP.
 AC AAF55630;
 XX
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DE Nucleotide sequence of a human endogenous retrovirus envelope protein.
 XX
 KW Envelope protein; HERV; syncytia formation; placental development;
 KW syncytia; cancer; cell adhesion; ss.
 XX
 OS Human endogenous retrovirus.
 XX
 XX Key Location/Qualifiers
 FH 762..2378
 CDS /*tag= a
 FT /product= "envelope protein"
 FT
 XX
 XX WO200116171-A1.
 XX
 XX
 PD 08-MAR-2001.
 XX
 PE 01-SEP-2000; 2000WO-FR002429.
 XX
 PR 01-SEP-1999; 99FR-00011141.
 PR 15-SEP-1999; 99FR-00011193.
 XX
 PA (INMR) BIO MERIEUX.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.

XX Mallet F, Cosset F, Blond J, Lavillette D, Bouton O, Ruggieri A;
 PI WPI: 2001-226676/23.
 DR P-PSDB; AAB67652.
 XX
 PT Detecting expression of human endogenous retrovirus envelope protein in
 PT cells of a tissue or culture, from its ability to induce syncytia.
 XX
 PS Disclosure; Page 44-45; 57pp; French.
 XX
 CC The present sequence encodes a human endogenous retrovirus envelope
 CC protein. The specification describes a method for detecting expression of
 CC an envelope protein from a human endogenous retrovirus (HERV), in cells,
 CC of a tissue or culture. The method comprises detecting syncytia formation
 CC due to the fusogenic properties of the envelope protein. Envelope
 CC polypeptides and polynucleotides are used to produce therapeutic or
 CC prophylactic compositions, particularly for treatment of cancer, to
 CC correct defects in placental development (or other natural formation of
 CC other types of syncytia), and to promote adhesion of cells in grafts or
 CC cellular repair processes. Expression of sequences antisense to the
 CC polynucleotide are used to prevent formation of syncytia
 XX
 XX Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 U; 0 Other;
 SQ
 Query Match 89.8%; Score 1330.6; DB 5; Length 2781;
 Best Local Similarity 93.7%; Pred. No. 0;
 Matches 1387; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATCATCTTTTCTTTTACTTGTCTCTTACCCCTTTGGCTCTCACT 60
 DB 762 ATGGCCCTCCCTTATCATATTTTCTTTTACTTGTCTTTTACCCCTTTTCACTCTCACT 821
 QY 61 GCACCCCTCCATGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTTATGAAGA 120
 DB 822 GCACCCCTCCATGCTGTGTATGACAGTAGCTCCCTTACCAAGAGTTTCTTATGAAGA 881
 QY 121 ACGCGGCTTCTTGGAATATTTGATGCCCCATCATATAGGAGTTTATCTAAGGGAATCTCC 180
 DB 882 ATGCAGGCTCCCGGAATATTTGATGCCCCATCGTATAGGAGTTTCTTCTAAGGGAATCTCC 941
 QY 181 ACCTTCACTGCCCAACCCATATGCCCCGCAACTGCTATTAACCTGCGACCTTTTGATG 240
 DB 942 ACCTTCACTGCCCAACCCATATGCCCCGCAACTGCTATTAACCTGCGACCTTTTGATG 1001
 QY 241 CATGCAAACTACTTATTTGAGAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT 300
 DB 1002 CATGCAAACTACTTATTTGAGAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT 1061
 QY 301 GGAGCCACTGCTGTTGGACTTACTTCAACCCATACCAAGTATGTCTGATGGGGTGGAAAT 360
 DB 1062 GGAGTCACTGCTGTTGGACTTACTTCAACCCATACCAAGTATGTCTGATGGGGTGGAGTT 1121
 QY 361 CAAGGTGAGCAAGAGAAAACAAAGTAAAGAGAGCAATCTCCCACTGAGCCCGGGACAT 420
 DB 1122 CAAGATCAGGCAAGAGAAAACAAAGTAAAGAGAGTATCTCCCACTGAGCCGGGTACAT 1181
 QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAGAACCTCCGTACC 480
 DB 1182 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAGAACCTCCGTACC 1241
 QY 481 CATACTCGCTGTTGAGCCTTATTAATACCAACCTCTCACTGGGTCTCCATGAGGTCTCAGCC 540
 DB 1242 CATACTCGCTGTTGAGCCTTATTAATACCAACCTCTCACTGGGTCTCCATGAGGTCTCAGCC 1301
 QY 541 CAAAACCTTACTAATGTTGATGTGCTCCCTCGCACTTCAGGCGCATATCATTTCAATC 600
 DB 1302 CAAAACCTTACTAATGTTGATGTGCTCCCTCGCACTTCAGGCGCATATCATTTCAATC 1361
 QY 601 CTTGTTCTGAAACAAATGGAACTTTCAGCAGAGAAATTAACCACTTCGTTTATGA 660
 DB 1362 CTTGTTCTGAAACAAATGGAACTTTCAGCAGAGAAATTAACCACTTCGTTTATGA 1421

QY 661 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCACCTGTGTAATAATTT 720
DB 1422 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCACCTGTGTAATAATTT 1481
QY 721 AGCAATACCTATAGACACACACACAGCTCCCAATGCATCAGGTGGTAAACCTCCCAACAGA 780
DB 1482 AGCAATACCTATAGACACACACACAGCTCCCAATGCATCAGGTGGTAACTCCTCCCAACAA 1541
QY 781 ATAGTCTGCTACCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
DB 1542 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 1601
QY 841 AATGCTCTTCAAGATCTATGTGCTTCTCTCATTTAGTGCCCTATGACCATCTAC 900
DB 1602 AATGCTCTTCAAGATCTATGTGCTTCTCTCATTTAGTGCCCTATGACCATCTAC 1661
QY 901 ACTGAACAAGATTTATACAAATCATGTCTGTAACCTTAAGCCCAACAAAGAGTACCCATT 960
DB 1662 ACTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATT 1721
QY 961 CTTCTCTTTTGTATCAGACAGAGTGTCTAGCAGACTAGGTACTGGCAATGGCAGTATC 1020
DB 1722 CTTCTCTTTTGTATAGGAGGGAGTGTCTAGTGCACTAGGTACTGGCAATGGCAGTATC 1781
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGCATGGAACAG 1080
DB 1782 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACG 1841
QY 1081 GTCACCTGACTCCCTGCTACCTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT 1140
DB 1842 GTCGCGGACTCCCTGCTACCTTTGCAAGATCAGCTTAACCTCCCTAGCAGCAGTAGTCCTT 1901
QY 1141 CAAATCGAAGAGCTTTAGACTGTCTAACCGCAAGAGGGGACCTGTTTATTTTA 1200
DB 1902 CAAATCGAAGAGCTTTAGACTGTCTAACCGCTGAAGAGGGGGAACCTGTTTATTTTA 1961
QY 1201 GGAGAAAGACGCTGTTTATTTATGTTAATCAATCAGAAATTTGTCAGTGAAGGTTAAAGAA 1260
DB 1962 GGGGAAGATGCTGTTTATTTATGTTAATCAATCGGAAATCGTCACTGAGAAAGTTAAAGAA 2021
QY 1261 ATTGAGAGTCGAATACAAATGTAGAGCAGAGAGCTTCAAAAACCGAACGCTGGGCGCTC 1320
DB 2022 ATTGAGAGTCGAATACAACTAGAGCAGAGAGCTTCAAAAACACTGGACCCCTGGGCGCTC 2081
QY 1321 CTCAGCAATGAGTGCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATG 1380
DB 2082 CTCAGCAATGAGTGCCTGGATTCTCCCTCTTAGGACCTCTAGCAGCTAATATATG 2141
QY 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTTTTAAAGTTTGTCTCTCCCAAGATT 1440
DB 2142 CTACTCTCTTTGGACCTGTATCTTTAACTCTTTTAACTTTGTTAACTTTGCTCTCCCAATC 2201
QY 1441 GAAGCTGTAAAGTCACAGATGGTCTTCAAAATGGAAACCCCA 1481
DB 2202 GAAGCTGTAAACTACAAATGGAGCCCAAGATCAGTCCCAA 2242

RESULT 19

AAA59215
ID AAA59215 standard; DNA; 7582 BP.

XX AAA59215;

XX 07-NOV-2000 (first entry)

XX Human endogenous retrovirus W (HERV-W) sequence.

XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX Human endogenous retrovirus.

XX FH

Key Location/Qualifiers

FT LTR
FT
FT LTR
FT LTR
FT primer_bind
FT CDS
FT CDS
FT CDS
FT misc_feature
FT LTR
FT polyA_signal
FT
PN WO200043521-A2.
PD 27-JUL-2000.
PF 21-JAN-2000; 2000WO-FR000144.
PR 21-JAN-1999; 99FR-00000888.
XX (INNR) BIO MERIEUX.
XX Paranhos-Baccala G, Mallet F, Voisset C;
XX WPI; 2000-49929/44.
XX
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
XX diagnosis of autoimmune disease and complications of pregnancy, contains
XX at least part of the gag gene.
XX
XX Disclosure; Page 49-52; 53pp; French.
XX
XX The present sequence represents an endogenous retrovirus, which is
XX associated with an autoimmune disease, and is integrated into the human
XX genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
XX HERV-W retrovirus is associated with autoimmune disease, failure of
XX pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
XX proteins derived from it, are useful for diagnosis of autoimmune disease
XX (specifically multiple sclerosis) and for monitoring pregnancy. The
XX nucleic acid fragments may also be used for in situ labelling of isolated
XX chromosomes, while the transcription product can be used to study or
XX monitor T cell proliferation in vitro
XX
XX Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;

Query Match 89.4%; Score 1324.2; DB 3; Length 7582;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 1365; Conservative 24; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATATCTTTCTTTACTGTTCTTACCCCTTTCTGCTCTCACT 60
DB 5581 ATGGSCCTCCCTTATCATATTTTCTCTASTGTTTTCCTTCTCACTCTCACT 5640
QY 61 GCACCCCTCCATGCTGCTGTACAAACAGTACCTCCCTTACCAAGAGTTCTATGAAGA 120
DB 5641 GCACCCCTCCATGCTGCTGTATGACAGTAGCTCCCTTACCMAGAGTTCTATGGAGA 5700
QY 121 ACGCGCTTCTCGGAATATTCATGCCCATCATATAGGATTATCTAAGGGAATCTC 180

XX Conrad B, Mach B;
XX WPI; 2001-316336/33.
XX P-PSDB; AAB75138.
XX New human retrovirus HERV-W ENV proteins/peptides having superantigen
XX activity useful for diagnosing and treating multiple sclerosis.
PS Claim 13; Fig 9; 94pp; English.
XX
XX On the basis of the PBS t-RNA motif used for the classification of human
XX endogenous retrovirus (HERVs) the full length endogenous provirus which
XX was located on the long arm of human chromosome 7 (7q21-22) has been
XX designated HERV-W. The present invention describes proteins or peptides
XX (I) having superantigen (SAG) activity comprising the ENV protein (ENV)
XX of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
XX have neuroprotective activity, and can be used in: vaccines; antisense-
XX therapy; and HERV-W SAG active-inhibitors. (I) and encoding DNA/RNA are
XX useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
XX disorders. (I) are also useful for identifying substances (and optionally
XX recovering) capable of binding to a retroviral superantigen associated
XX with MS, substances capable of blocking SAG activity and substances
XX capable of blocking transcription or translation of HERV-W retroviral
XX superantigen. A protein or peptide derived from (I), modified to be
XX devoid of SAG activity and being capable of generating an immune response
XX against HERV-W retroviral SAG is useful in therapy. Nucleic acid
XX molecules encoding (I) are useful as vaccines against MS. Substances
XX capable of blocking SAG activity, capable of binding to a retroviral
XX superantigen associated with MS, or capable of blocking transcription or
XX translation of HERV-W retroviral superantigen for use in treating or
XX preventing MS, obtained using (I) are useful for the treatment and
XX prevention of MS. (I) and nucleic acids encoding them are useful for
XX diagnosing autoimmune disease. The present sequence encodes the
XX specifically claimed envelope protein of HERV-W designated G
XX
XX Sequence 2782 BP; 741 A; 768 C; 564 G; 709 T; 0 U; 0 Other;
Query Match 88.3%; Score 1308.2; DB 5; Length 2782;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1373; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 1 ATGCGCCCTCCCTTATCATCTTTCTCTTACTGCTCTTACCCCTTTGCGTCTCACT 60
DB 763 ATGCGCCCTCCCTTATCATCTTTCTCTGAGTGTCTTTCACCCCTGTTTCACTCTCACT 822
QY 61 GCACCCCTCCATGCTGTGATCAACCCAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 823 GCACCCCTCCATGCGCTGTATGACCCAGTAGCTCCCTCACCAGAGTTTCTATGGAGA 882
QY 121 ACGCGGCTTCTGGAATATTGATGCGCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180
DB 883 ATGACGGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTCTAAGGGAAACCC 942
QY 181 ACCTTCACTGCGCCACCATATGCGCGCACTGCTATTAATCTGACCTTTTGTGATG 240
DB 943 ACCTTCACTGCGCCACCATATGCGCGCACTGCTATTAATCTGACCTTTTGTGATG 1002
QY 241 CATGCAATATCTATTATTTGACAGGGGAAATGATTAATCTAGTTGTCTGAGGAGCTT 300
DB 1003 CATGCAATATCTATTATTTGACAGGGGAAATGATTAATCTAGTTGTCTGAGGAGCTT 1062
QY 301 GGAGCCACTGTCTGTTGACTTACTTCCACCATACAGATATGCTGATGGGGTGGAAATT 360
DB 1063 GGAGTCACTGTCTGTTGACTTACTTCCACCATACAGATATGCTGATGGGGTGGAGTT 1122
QY 361 CAAGGTGAGGCAAGGAAACAACTAAGGAAGCAATCTCCCACTGACCCCGGGGACAT 420
DB 1123 CAAGGTGAGGCAAGGAAACAACTAAGGAAGCAATCTCCCACTGACCCCGGGGACAT 1182
QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
DB 1183 GGCACCTTAGCCCTCAAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 1242

QY 481 CATACTCGCTGCTGAGCCTATTATATACACCCCTCACTCGGCTCATGAGGCTCTCAGCC 540
DB 1243 CATACTCGCTGCTGAGCCTATTATATACACCCCTCACTCGGCTCTCAGGCTCTCGGCC 1302
QY 541 CAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGCTCACTTCAGGGCATACATTTCAATC 600
DB 1303 CAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGAACTTCAGGGCATATGTTTCAATC 1362
QY 601 CTTGTTCTCTGAACAATGGAACCAACTTCAGCACAGAAATAAACACCACTTCCTCGTTTGTAGTA 660
DB 1363 CTTGTTCTCTGAACAATGGAACCAACTTCAGCACAGAAATAAACACCACTTCCTCGTTTGTAGTA 1422
QY 661 GGACCTCTTGTGTTTCCAACTCTGGAATAAACCACTTACCTCAAACTCACTGTGTAAATTT 720
DB 1423 GGACCTCTTGTGTTTCCAACTCTGGAATAAACCACTTACCTCAAACTCACTGTGTAAATTT 1482
QY 721 AGCAATCTATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACAGA 780
DB 1483 AGCAATCTATACATACACACACCACTCCCAATGATCAGGTGGGTAACTCTCCACACAAA 1542
QY 781 ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTTG 840
DB 1543 ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTTG 1602
QY 841 AATGGCTCTTCAAGAACTATGCTCTCTCTCAATTTCTTAGTGCCCTTATGACCATCTAC 900
DB 1603 AATGGCTCTTCAAGAACTATGCTCTCTCTCAATTTCTTAGTGCCCTTATGACCATCTAC 1662
QY 901 ACTGAAACAAGATTTATCAATCATGCTGTAACCTTAAGCCCAACAAACAAAGATACCCATT 960
DB 1663 ACTGAAACAAGATTTATACAGTTATGTATCATATCTAAGCCCGCAACAAACAAAGATACCCATT 1722
QY 961 CTTCTCTTTGTTATCAGACGAGGAGTGCTAGGCAGACTAGTACTGTCATTTGGCAGTATC 1020
DB 1723 CTTCTCTTTGTTATAGGACGAGGAGTGCTAGGTGCTAGTACTGTCATTTGGCAGTATC 1782
QY 1021 ACAACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGGTGATGGAACAG 1080
DB 1783 ACAACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGGGACATGGAACAG 1842
QY 1081 GTCACCTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1843 GTCGCGGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1902
QY 1141 CAAATCGAAGAGCTTTAGAGCTTTAGAGCTTTAGAGCTTTAGAGCTTTAGAGCTTTAGAGCT 1200
DB 1903 CGAAATCGAAGAGCTTTAGAGCTTTAGAGCTTTAGAGCTTTAGAGCTTTAGAGCTTT 1962
QY 1201 GGAGAAAGACGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1260
DB 1963 GGAGAAAGACGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2022
QY 1261 ATTTCAGATCGAATATCAATGTAGACGAGAGAGCTTCAAAACACCGAAGCCGTGGGGCTC 1320
DB 2023 ATTTCAGATCGAATATCAATGTAGACGAGAGAGCTTCAAAACACCGAAGCCGTGGGGCTC 2082
QY 1321 CTAGCCCAATGGATGCGCTGCGGTCTCCCTCTCTTAGAGCTCTTAGAGCTCTTAATATTG 1380
DB 2083 CTAGCCCAATGGATGCGCTGCGGTCTCCCTCTCTTAGAGCTCTTAGAGCTCTTAATATTG 2142
QY 1381 TTACTCTCTCTTTGGAGCCCTGTATCTTTTAACTCTCTTCTTAAAGTTTGTCTCTTCCAGAA 1440
DB 2143 CTACTCTCTCTTTGGAGCCCTGTATCTTTTAACTCTCTTCTTAAAGTTTGTCTCTTCCAGAA 2202
QY 1441 GAAGCTGTAAAGCTTACAGATGGTCTTTTAAATAATGGAACCCCA 1481
DB 2203 GAAGCTGTAAAGCTTACAGATGGTCTTTTAAATAATGGAACCCCA 2243

RESULT 22
AA25661
ID AAX25661 standard; cDNA to mRNA; 2782 BP.

XX AC AAX25661; Db 1123 CAAGATCAGGCAAGAGAAAACATGTAAGAAGTAATCTCCCAACTCACCGGGTACAT 1182
XX DT 21-MAY-1999 (first entry) Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAAAAATACATAGAAACCTCGGTACC 480
XX DE Human endogenous retrovirus W clone cl.PH74. Db 1183 GGCACCTCTAGCCCTACAAAGGACTAGTCTCTCAAAAATACATAGAAACCTCGGTACC 1242
XX KW Clone; human endogenous retrovirus; genome; autoimmune disease; Qy 481 CATACTGCCTGGTGGAGCTATTTAATACCAACCTCTACTCGGCTCCATAGAGTCTCAGCC 540
XX KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes; Db 1243 CATACTGCCTGGTGGAGCTATTTAATACCAACCTCTACTCGGCTCCATAGAGTCTCAGCC 1302
XX KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss. Qy 541 CAAAACCTCTACTAATCTGTGGATGTGCTCTCCCTCGCATCTTCAGGCCCATACATTTCAATC 600
XX OS Human endogenous retrovirus. Db 1303 CAAAACCTCTACTAATCTGTGGATGTGCTCTCCCTCGCATCTTCAGGCCCATATGTTTCAATC 1362
XX PN WO902696-A1. Qy 601 CCTGTTCTGAACAAATGGAAACAATCTCAGCAGAGAAATAACACACACTTCCTGTTTAGTA 660
XX PD 21-JAN-1999. Db 1363 CCTGTACTGAACAAATGGAAACAATCTCAGCAGAGAAATAACACACACTTCCTGTTTAGTA 1422
XX PF 06-JUL-1998; 98WO-FR001442. Qy 661 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAAACCTCACCTGTGTAAATTT 720
XX PR 07-JUL-1997; 97FR-00008815. Db 1423 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAAACCTCACCTGTGTAAATTT 1482
XX PA (INMR) BIO MERIEUX. Qy 721 AGCAATCTATAGACACAAACAGCTCCCAATGCAATGCAATGAGTGGGTAAACCTCCCAACAGA 780
XX PI Besene F, Blond J, Bouton O, Mandrand B, Mallet F; Db 1483 AGCAATCTATAGACACAAACAGCTCCCAATGCAATGCAATGAGTGGGTAAACCTCCCAACAGA 1542
XX DR WPI; 1999-120897/10. Qy 781 ATAGTCTGCCCTACCCCTCAGGAATATTTTGTGTGTGTAGTCTCAGCCTATCATTTGTTG 840
XX XX New nucleic acid sequences from human endogenous retrovirus-W - expressed Pt exclusively in placenta and useful in diagnosis and therapy of autoimmune disease, and abnormal or failed pregnancy. Db 1543 ATAGTCTGCCCTACCCCTCAGGAATATTTTGTGTGTGTAGTCTCAGCCTATCATTTGTTG 1602
XX PS Claim 1; Page 60-63; 106pp; French. Qy 841 AATGGCTCTCAGAAATCTATGTGCTTCTCTCATTTAGTGCCCTCTATGACCATCTAC 900
XX SS This sequence represents clone cl.PH74 of the human endogenous retrovirus (HERV) W genome. The nucleic acids, their fragments or peptides encoded by them are markers of autoimmune disease (e.g. multiple sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-dependent diabetes and related pathologies) and of abnormal or unsuccessful pregnancy and can be used as chromosomal markers for susceptibility to these conditions, or proximity markers of genes associated with this susceptibility. Qy 1603 AATGGCTCTCAGAAATCTATGTGCTTCTCTCATTTAGTGCCCTCTATGACCATCTAC 1662
XX SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other; Qy 901 ACTGAACAGATTTTATACAATCATGTCTAAGCCCTCAACAAAGAGTACCCATT 960
Qy 1663 ACTGAACAGATTTTATACAATCATGTCTAAGCCCTCAACAAAGAGTACCCATT 1722
Qy 961 CTTCCCTTTGTTATCAGAGCAGGAGTGTAGGACAGACTAGGTACTGGCATGGCAGTATC 1020
Db 1723 CTTCCCTTTGTTATAGGAGCAGGAGTGTAGTGTGCACTAGGTACTGGCATGGCAGTATC 1782
Qy 1021 ACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGGTGACATGAACAG 1080
Db 1783 ACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGGTGACATGAACAG 1842
Qy 1081 GTCAGTACTCCTCGTCACTTGTGCAAGATCAACTTAACCTCCTAGCAGCAGTACTCTTT 1140
Db 1843 GTCGCCGACTCCTCGTCACTTGTGCAAGATCAACTTAACCTCCTAGCAGCAGTACTCTTT 1902
Qy 1141 CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAAACCTGTTTATTTTAA 1200
Db 1903 CAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGGGAAACCTGTTTATTTTAA 1962
Qy 1201 GGAGAGAAACGCTGTTTATTTATTTAATCAATCAGAAATGTCATGAGAAAGTTAAAGAA 1260
Db 1963 GGGGAGAAATGCTGTTTATTTAATCAATCGGAAATCGTCACTGAGAAAGTTGAGAA 2022
Qy 1261 ATTCCAGATCGAATAAATGTTAGAGCAGAGAGCTTCAAAACACGAAACGCTGGGSCCTC 1320
Db 2023 ATTCCAGATCGAATAAATGTTAGAGCAGAGAGCTTCGAAACACTGGACCTGGGSCCTC 2082
Qy 1321 CTCAGCAATGATGATGCCCTGGGTCTCCCTCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1380
Db 2083 CTCAGCGATGATGATGCCCTGGATCTCCCTCTTCTTAGGACCTCTAGCAGCTATAATTG 2142
Qy 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGAAAT 1440
Db 2143 CTACTCCTCTTTGGACCCCTGTATCTTTGACCTCTTTGACCTCTTTGTTAACTTTGTCTCTTCCAGAAATC 2202
Qy 1441 GAAGCTGTAAAGCTTACAGATGCTTACAAATGGAAACCCCA 1481

XX AC AAX25661; Db 1123 CAAGATCAGGCAAGAGAAAACATGTAAGAAGTAATCTCCCAACTCACCGGGTACAT 1182
XX DT 21-MAY-1999 (first entry) Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAAAAATACATAGAAACCTCGGTACC 480
XX DE Human endogenous retrovirus W clone cl.PH74. Db 1183 GGCACCTCTAGCCCTACAAAGGACTAGTCTCTCAAAAATACATAGAAACCTCGGTACC 1242
XX KW Clone; human endogenous retrovirus; genome; autoimmune disease; Qy 481 CATACTGCCTGGTGGAGCTATTTAATACCAACCTCTACTCGGCTCCATAGAGTCTCAGCC 540
XX KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes; Db 1243 CATACTGCCTGGTGGAGCTATTTAATACCAACCTCTACTCGGCTCCATAGAGTCTCAGCC 1302
XX KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss. Qy 541 CAAAACCTCTACTAATCTGTGGATGTGCTCTCCCTCGCATCTTCAGGCCCATACATTTCAATC 600
XX OS Human endogenous retrovirus. Db 1303 CAAAACCTCTACTAATCTGTGGATGTGCTCTCCCTCGCATCTTCAGGCCCATATGTTTCAATC 1362
XX PN WO902696-A1. Qy 601 CCTGTTCTGAACAAATGGAAACAATCTCAGCAGAGAAATAACACACACTTCCTGTTTAGTA 660
XX PD 21-JAN-1999. Db 1363 CCTGTACTGAACAAATGGAAACAATCTCAGCAGAGAAATAACACACACTTCCTGTTTAGTA 1422
XX PF 06-JUL-1998; 98WO-FR001442. Qy 661 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAAACCTCACCTGTGTAAATTT 720
XX PR 07-JUL-1997; 97FR-00008815. Db 1423 GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAAACCTCACCTGTGTAAATTT 1482
XX PA (INMR) BIO MERIEUX. Qy 721 AGCAATCTATAGACACAAACAGCTCCCAATGCAATGCAATGAGTGGGTAAACCTCCCAACAGA 780
XX PI Besene F, Blond J, Bouton O, Mandrand B, Mallet F; Db 1483 AGCAATCTATAGACACAAACAGCTCCCAATGCAATGCAATGAGTGGGTAAACCTCCCAACAGA 1542
XX DR WPI; 1999-120897/10. Qy 781 ATAGTCTGCCCTACCCCTCAGGAATATTTTGTGTGTGTAGTCTCAGCCTATCATTTGTTG 840
XX XX New nucleic acid sequences from human endogenous retrovirus-W - expressed Pt exclusively in placenta and useful in diagnosis and therapy of autoimmune disease, and abnormal or failed pregnancy. Db 1543 ATAGTCTGCCCTACCCCTCAGGAATATTTTGTGTGTGTAGTCTCAGCCTATCATTTGTTG 1602
XX PS Claim 1; Page 60-63; 106pp; French. Qy 841 AATGGCTCTCAGAAATCTATGTGCTTCTCTCATTTAGTGCCCTCTATGACCATCTAC 900
XX SS This sequence represents clone cl.PH74 of the human endogenous retrovirus (HERV) W genome. The nucleic acids, their fragments or peptides encoded by them are markers of autoimmune disease (e.g. multiple sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-dependent diabetes and related pathologies) and of abnormal or unsuccessful pregnancy and can be used as chromosomal markers for susceptibility to these conditions, or proximity markers of genes associated with this susceptibility. Qy 1603 AATGGCTCTCAGAAATCTATGTGCTTCTCTCATTTAGTGCCCTCTATGACCATCTAC 1662
XX SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other; Qy 901 ACTGAACAGATTTTATACAATCATGTCTAAGCCCTCAACAAAGAGTACCCATT 960
Qy 1663 ACTGAACAGATTTTATACAATCATGTCTAAGCCCTCAACAAAGAGTACCCATT 1722
Qy 961 CTTCCCTTTGTTATCAGAGCAGGAGTGTAGGACAGACTAGGTACTGGCATGGCAGTATC 1020
Db 1723 CTTCCCTTTGTTATAGGAGCAGGAGTGTAGTGTGCACTAGGTACTGGCATGGCAGTATC 1782
Qy 1021 ACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGGTGACATGAACAG 1080
Db 1783 ACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGGTGACATGAACAG 1842
Qy 1081 GTCAGTACTCCTCGTCACTTGTGCAAGATCAACTTAACCTCCTAGCAGCAGTACTCTTT 1140
Db 1843 GTCGCCGACTCCTCGTCACTTGTGCAAGATCAACTTAACCTCCTAGCAGCAGTACTCTTT 1902
Qy 1141 CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAAACCTGTTTATTTTAA 1200
Db 1903 CAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGGGAAACCTGTTTATTTTAA 1962
Qy 1201 GGAGAGAAACGCTGTTTATTTAATCAATCAGAAATGTCATGAGAAAGTTAAAGAA 1260
Db 1963 GGGGAGAAATGCTGTTTATTTAATCAATCGGAAATCGTCACTGAGAAAGTTGAGAA 2022
Qy 1261 ATTCCAGATCGAATAAATGTTAGAGCAGAGAGCTTCAAAACACGAAACGCTGGGSCCTC 1320
Db 2023 ATTCCAGATCGAATAAATGTTAGAGCAGAGAGCTTCGAAACACTGGACCTGGGSCCTC 2082
Qy 1321 CTCAGCAATGATGATGCCCTGGGTCTCCCTCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1380
Db 2083 CTCAGCGATGATGATGCCCTGGATCTCCCTCTTCTTAGGACCTCTAGCAGCTATAATTG 2142
Qy 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGAAAT 1440
Db 2143 CTACTCCTCTTTGGACCCCTGTATCTTTGACCTCTTTGACCTCTTTGTTAACTTTGTCTCTTCCAGAAATC 2202
Qy 1441 GAAGCTGTAAAGCTTACAGATGCTTACAAATGGAAACCCCA 1481

Db	2083	CTCAGCGGATGGATGCGCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATTTG	2144			
Qy	1381	TTACTCTCTTTTGGAGCCCTGTATCTTTTAACCTCTCTTTTAAAGTTTGTCTTCCAGAA	1440			
Db	2143	CTACTCTCTTTTGGAGCCCTGTATCTTTTGACCTCTCTTTTAACTTTGTCTCTTCCAGATC	2202			
Qy	1441	GAGCTGTAAAGCTACAGATGGTCTTACAATATGGAACCCCA	1481			
Db	2203	GAAAGCTGTGAAACTACAAATGGAGCCCAAGATGCAAGTCCAA	2243			
RESULT 24						
ADE09587						
ID	ADE09587 standard; DNA; 3464 BP.					
XX	ADE09587;					
XX	29-JAN-2004 (first entry)					
DE	Novel DNA-related contig nucleotide sequence #309.					
XX	novel gene; novel protein; tissue marker; molecular weight marker;					
KW	chromosome marker; genetic disorder; contig; ds.					
XX	Unidentified.					
OS						
XX	WO2003054152-A2.					
FN						
XX	03-JUL-2003.					
PD						
XX	10-DEC-2002; 2002WO-US039555.					
PF						
XX	10-DEC-2001; 2001US-0339739P.					
PR	11-DEC-2001; 2001US-0339453P.					
PR	14-MAR-2002; 2002US-0365091P.					
PR	14-MAR-2002; 2002US-0365384P.					
PR	12-APR-2002; 2002US-0372381P.					
PR	12-APR-2002; 2002US-0372615P.					
PR	22-APR-2002; 2002US-00128558.					
PR	24-APR-2002; 2002US-0376045P.					
XX						
PA	(HYSE-) HYSEQ INC.					
XX						
PI	Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;					
PI	Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;					
PI	Ma Y, Wang D, Chen R, Xu C, Boyle BJ;					
XX						
DR	WPI; 2003-569235/53.					
XX						
PT	New polynucleotides, useful for expressing recombinant proteins for					
PT	analysis, characterization or therapeutic use, or as markers for tissues					
PT	in which the corresponding protein is preferentially expressed.					
XX						
PS	Disclosure; SEQ ID NO 2131; 1177pp; English.					
XX						
CC	The invention comprises the amino acid and coding sequences of novel					
CC	proteins. The DNA and protein sequences of the invention are useful as:					
CC	markers for tissues in which the corresponding protein is preferentially					
CC	expressed; as molecular weight markers on gels; as chromosome markers or					
CC	tags; to identify chromosomes or to map related gene positions; and to					
CC	compare with endogenous DNA sequences in patients to identify potential					
CC	genetic disorders. The present DNA sequence was used in the					
XX	exemplification of the invention.					
XX						
SQ	Sequence 3464 BP; 1038 A; 913 C; 666 G; 845 T; 0 U; 2 Other;					
Query Match 87.7%; Score 1298.2; DB 10; Length 3464;						
Best Local Similarity 93.0%; Pred. No. 0;						
Matches 1383; Conservative 0; Mismatches 98; Indels 6; Gaps 2;						
Qy	1 ATGGCCCTCCCTTATCATACTTTCTTTACTGTCTTACCCCTTCGCTCTCACT 60					
Db	1430 ATGGCCCTCCCTTATCATATTTCTTTACTGTCTTACCTCTTCACTCTCACT 1489					

1135 GTCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAGAGGGGAACCTGTGTTA 1194
 2570 GTCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGGGGAACCTGTGTTA 2629
 1195 TTTTGGAGAGAGCGCTGTTATATGTTAATCAATCAATCGAATTTGCTACTGAGAAGTT 1254
 2630 CTTTGGAGAGAGAGCGCTGTTATATGTTAATCAATCGAATTTGCTACTGAGAAGTT 2689
 1255 AAGAAATTCGAGATCGAATACAAATGAGAGAGAGGCTTCAAAACACCGGAACGCTGG 1314
 2690 AAGAAATTCGAGATCGAATACAAATGAGAGAGAGGCTTCAAAACACCGGAACGCTGG 2749
 1315 GGCTCTCTCAGCAATGAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
 2750 GGCTCTCTCAGCAATGAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2809
 1375 ATATGTTACTCTCTTGTGAGCGCTGATCTTTAACTCTGCTGCTGCTGCTGCTGCTGCT 1434
 2810 ATATGTTACTCTCTTGTGAGCGCTGATCTTTAACTCTGCTGCTGCTGCTGCTGCTGCT 2869
 1435 AGAATTGAAGCTGTAAGCTACAGATGCTTACAAATGGAACCCCA 1481
 2870 AGAATGAAGCTGTAAGCTACAGATGCTTACAAATGGAACCCCA 2916

RESULT 25

ADFS9718 ID ADFS9718 standard; cDNA; 9502 BP.

AC ADFS9718;

XX ADFS9718;

XX 12-FEB-2004 (first entry)

XX Human contig polynucleotide sequence SEQ ID NO:2085.

XX biological activity; genetic engineering; hybridisation probe; oligomer;

XX primer; chromosome mapping; gene mapping; recombinant protein production;

XX human; gene; ss.

XX Homo sapiens.

XX WO2003080795-A2.

XX 02-OCT-2003.

XX 09-AUG-2002; 2002WO-US025485.

XX 09-AUG-2001; 2001US-0311261P.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX WPI; 2003-876918/81.

XX P-PSDB; ADF60170.

XX New polynucleotides, useful as hybridization probes, oligomers or

XX primers, for chromosome or gene mapping, for the recombinant production

XX of proteins, and for generating antisense DNA or RNA.

XX Example 2; SEQ ID NO 2085; 571pp; English.

XX The present invention describes isolated polynucleotide sequences (I),

XX which encode polypeptides (II) with biological activity. Also described:

XX (1) a vector comprising (I); (2) an expression vector comprising (I); (3)

XX a host cell genetically engineered to comprise (I) which is operatively

XX associated with a regulatory sequence that modulates expression of (I) in

CC (I). The polynucleotides (I) can be used as hybridisation probes,
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant
 CC production of proteins, and for generating antisense DNA or RNA. The
 CC present sequence represents a human contig polynucleotide sequence, which
 CC is used in an example from the present invention.

XX
 SQ Sequence 9502 BP; 2813 A; 2433 C; 1992 G; 2263 T; 0 U; 1 Other;

Query Match 87.7%; Score 1298.2; DB 10; Length 9502;
 Best Local Similarity 93.0%; Pred. No. 0;
 Matches 1383; Conservative 0; Mismatches 98; Indels 6; Gaps 2;

QY 1 ATGGCCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTGGTCTCTCACT 60
 DB 1430 ATGGCCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1489

QY 61 GCACCCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 120
 DB 1490 GCACCCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1549

QY 121 ACCTGGCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 180
 DB 1550 ATGCAGGCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1609

QY 181 ACCTTCACCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 240
 DB 1610 ACCTTCACCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1669

QY 241 CATGCAATCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 300
 DB 1670 CATGCAATCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1729

QY 301 GGAGCCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 360
 DB 1730 GGAGCTCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1789

QY 361 CAAGGTGAGCAAG 420
 DB 1790 CAAGATCAGCAAG 1849

QY 421 AGCACCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 480
 DB 1850 AGCACCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1909

QY 481 CATACTCGCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 540
 DB 1910 CATACTCGCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 1969

QY 541 CAAACCCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 600
 DB 1970 CAAACCCCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 2029

QY 601 CTTGTTCTCTGAAATGGAACCACTTTCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 2030 CTTGTTCTCTGAAATGGAACCACTTTCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2089

QY 661 GGACCTCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 720
 DB 2090 GGACCTCTCTCTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTTCACTCTCTCACT 2149

QY 721 AGCAATCTATAGACACCAACCTCTCCCAATGATGAGGTGGGTAAACCTCTCCACACGGA 780
 DB 2150 AGCAATCTATAGACACCAACCTCTCCCAATGATGAGGTGGGTAAACCTCTCCACACGGA 2209

QY 781 ATAGTCTGCT 840
 DB 2210 ATAGTCTGCT 2269

QY 841 AATGGCT 900
 DB 2270 AATGGCT 2329

QY 901 ACTGAACAAGATTTTATCAATCATGTGTC----GTACCTAAAGCCCCCAACAAAGAGTACC 956

Db 2330 ACTGAACAGATTATACAGTTATGTCAATATCTTAAGCCCGCCGACAAAGAGTACC 2389
 QY 957 CATCTCTCTTTTGTATCAGACGAGGAGTGTAGGACACTAGGTACTGGCAATTGGCAG 1016
 Db 2390 CATCTCTCTTTTGTATTTGGAGCAGGAGTACTAGGCGGACTAGTACTGGCAATTGGCGG 2449
 QY 1017 TATCACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATTAATGGTGACATGGA 1076
 Db 2450 TATTACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATTAATGGTGACATGGA 2509
 QY 1077 ACAGGT--CACTGACTCCCTGTGCTACCTTGCAGATCACTTTAACTCCCTAGCAGAGTA 1134
 Db 2510 ACAGGTGCGCCGACTCCCTGTGCTACCTTGCAGATCACTTTAACTCCCTAGCAGAGTA 2569
 QY 1135 GTCCTTCAAAATCGAAGAGCTTTAGACTTTGCTTAAACCGCCAAAGAGGGGGAACCTGTTTA 1194
 Db 2570 GTCTTTCAAAATCGAAGAGCTTTAGACTTTGCTTAAACCGCTGAAAGAGGGGGAACCTGTTTA 2629
 QY 1195 TTTTAGGAGAAAGCGCTGTATATGTTATCAATCCAGAAATGTCACCTGAGAAAGTT 1254
 Db 2630 CTTTTAGGGGAAGATGCTGTGTACTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTT 2689
 QY 1255 AAGAAATTCGAGATCGAATCAATCTAGACGAGGAGCTTCAAAACACCGAAGCGCTGG 1314
 Db 2690 AAGAAATTCGAGATCGAATCAACGTAGACGAGGAGCTTCAAAACACCTGACCTGG 2749
 QY 1315 GGCCTCCTCAGCAATGGAATGAGCTGCTGGTCTCCCTTCTTAGGAGCTCTAGCAGCTCTA 1374
 Db 2750 GGCCTCCTCAGCAATGGAATGAGCTGCTGGTCTCCCTTCTTAGGAGCTCTAGCAGCTATA 2809
 QY 1375 ATATTGTTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCC 1434
 Db 2810 ATATTGTTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCC 2869
 QY 1435 AGAATCGAAGCTGTAAGAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 Db 2870 AGAATCGAAGCTGTAAGAGCTACAAATGGAGCCCAAGATGCAGTCCAA 2916

RESULT 26

ID ABN97978 standard; DNA; 46340 BP.
 XX AC ABN97978;
 XX DT 01-AUG-2002 (first entry)
 XX DE Human retroviral sequence H13.
 XX KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
 XX OS multiple sclerosis; ds.
 XX OS Human endogenous retrovirus.
 XX PN WO9667395-A1.
 XX PD 29-DEC-1999.
 XX PF 23-JUN-1999; 99WO-FR001513.
 XX PR 23-JUN-1998; 98PR-00007920.
 XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX PI Alliel PM, Perin J, Rieger F;
 XX WPI; 2000-160587/14.
 XX PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
 XX PT for diagnosis, treatment and prevention of autoimmune and neurological
 XX PT diseases.

PS Claim 15; Page 186-199; 225pp; French.
 XX The present invention relates to new nucleic acid sequences of human
 CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
 CC Regulatory elements associated with HERV-7q may alter expression of other
 CC genes (even remote genes) on the same chromosome, inducing immunological
 CC and/or neurological changes (which may be pathological or protective/
 CC curative). HERV-7q peptides can be used to improve efficiency of the
 CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
 CC sequences can be used in immunogenic or vaccinating compositions, for
 CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention
 XX SQ Sequence 46340 BP; 16104 A; 8738 C; 8434 G; 13064 T; 0 U; 0 Other;
 Query Match 86.6%; Score 1282; DB 3; Length 46340;
 Best Local Similarity 92.3%; Pred. No. 0;
 Matches 1395; Conservative 0; Mismatches 85; Indels 32; Gaps 3;
 QY 1 ATGGCCCTCCCTTATCATACTTTCTTTACTGTTCTTACCCCTTTTGGCTCTCACT 60
 Db 31855 ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTACCCCTTTTCACTCTCACT 31914
 QY 61 GCACCCCTCCATGCTGTACAAACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
 Db 31915 GCACCCCTCCATGCTGTACAAACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 31974
 QY 121 ACGCGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
 Db 31975 ATGGCGCTTCCAGAGAAATATTGATGCCCATCAATAGGAGTTTACCTTAAAGGAAACTCC 32034
 QY 181 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATACTGCCCACTCTTTGCGATG 240
 Db 32035 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATACTGCCCACTCTTTGCGATG 32094
 QY 241 CATGCAAAATCTATTATTGGACAGGAAATGATTAACTCTAGTTGCTCGAGAGACTT 300
 Db 32095 CATGCAAAATCTATTATTGGACAGGAAATGATTAACTCTAGTTGCTCGAGAGACTT 32154
 QY 301 GGAGCCACTGCTGTGTGGACTTACTTTACCCCATACAGTATGTTCTGTATGGGGTGGATT 360
 Db 32155 GGAGCCACTGCTGTGTGGACTTACTTTACCCCATACAGTATGTTCTGTATGGGGTGGATT 32214
 QY 361 CAGGCTCAGGCAAGAGAAACCAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
 Db 32215 CAGGCTCAGGCAAGAGAAACCAAGTAAAGGAAGTAACTCCCACTGACCCGGGTACAT 32274
 QY 421 AGCACCCCTTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCTTCGGTACC 480
 Db 32275 AGCACCCCTTAGCCCTTACAAAGGACTAGATCTCTTAAAGGACTACATGAAACCTTCATACC 32334
 QY 481 CATACTGCGCTGGTGAGCCTATTATTAATACCCCTTCACTCGGCTCCATGAGGTCTCAGCC 540
 Db 32335 CATACTGCGCTGGTGAGCCTATTATTAATACCCCTTCACTCGGCTCCATGAGGTCTCAGCC 32394
 QY 541 CAAAACCTTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCGCATACATTTTCAATC 600
 Db 32395 CAAAACCTTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCGCATGATTTTCAATC 32454
 QY 601 CCTGTTTCTGAAACAATGGAACTTTCAGCACAGAAATAAAACCACTTCCGTTTTAGTA 660
 Db 32455 CCTGTACTGTAACAATGGAACTTTCAGCACAGAAATAAAACCACTTCCGTTTTAGTA 32514
 QY 661 GGACCTCTTGTTCNAATCTGGAATAACCCATCTCAACCTCAACCTGTGTAAATTT 720
 Db 32515 GGACCTCTTGTTCNAATCTGGAATAACCCATCTCAACCTCAACCTGTGTGTAAATTT 32574
 QY 721 AGCAATCTATAGACAAACAGCTCCCAATGATCATCAGGTGGGTAAACACCTCCACACGA 780
 Db 32575 AGCAATCTATAGACAAACAGCTCCCAATGATCATCAGGTGGGTAACTCCTCCACACGA 32634

Db 1222 CAAACCCCTACTACTGTTGTTGTCCTCCCTGTAATTCAGGCCATGCAATTCATC 1163
Qy 601 CTGTTCTGGAACAATGGAACAATTCAGCAAGAAATAAACAACATTCCTGTTTAGTA 660
Db 1162 CCGTACCTGGAACAATGGAACAATTCAGCAAGAAATAAACAACATTCCTGTTTAGTA 1103
Qy 661 GCACCTCTGTTTCCAAATCTGGAATAAACCCTACCTCAACCTCAGCTGTGTAATTT 720
Db 1102 GCACCTCTGTTTCCAAATCTGGAATAAACCCTACCTCAACCTCAGCTGTGTAATTT 1043
Qy 721 AGCAATACCTATAGACACAAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
Db 1042 AGCAATACCTATAGACACAAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 983
Qy 781 ATAGTCTGCTACCTCCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATGTTTG 840
Db 982 ATAGTCTGCTACCTCCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATGTTTG 923
Qy 841 AATGGCTCTCAGAACTATGCTTCTCTCTCATCTTAGTCCCTATGACCTATGACCATCTAC 900
Db 922 AATGGCTCTCAGAACTATGCTTCTCTCTCATCTTAGTCCCTCTATGACCTATGACCATCTAC 864
Qy 901 ACTGAACAAGATTTATACAAATCATGCTGTAACCTTAAGCCCCCAACAACAAAGAGTACCCATT 960
Db 863 ACTGAACAAGATTTATACAAATCATGCTGTAACCTTAAGCCCCCAACAACAAAGAGTACCTATT 804
Qy 961 CTTCTCTTTGTTATACAGACGAGGTGCTAGGACACTAGGTACTGGCAATGGGAGTATC 1020
Db 803 CTTCTCTTTGTTATCGAGACGAGGTGCTAGGTGGAAGTCTGGCAATGGGAGTATC 744
Qy 1021 ACACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATAAATGGTGCACATGGAACAG 1080
Db 743 ACACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATAAATGGTGCACATGGAATGG 684
Qy 1081 GTCACTGACTCCCTGCTGCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Db 683 GTTGCCGACTCCCTGCTGCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 624
Qy 1141 CAAATCGAAGAGCTTTAGACTGCTTAACCGCCAAAGAGGGGAACTGTTATTTTAA 1200
Db 623 CAAATCGAAGAGCTTTAGACTGCTTAACCGCCAAAGAGGGGAACTGTTATTTTAA 564
Qy 1201 GGAGAGAGAGCTG-----TTATTATGTTAATCAA 1230
Db 563 GGGGAAGAATGTTGTTATGTTATTTAGGGGAAGAATGTTGTTATGTTAATCAA 504
Qy 1231 TCAGAAATTTGCTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAG 1290
Db 503 TCCTGAATTTGCTCACAGAGAAAGTTGAAGAAATTCGAGATCGAATACAAGTAGACAGAG 444
Qy 1291 GAGCTTC-AAAACACGAGCGTGGGGCTCCTCAGCAGCAATGATGCGCTGGTCTCC 1349
Db 443 GAGCTTC-AAAACACGAGCGTGGGGCTCCTCAGCAGCAATGATGCGCTGGTCTCC 384
Qy 1350 CTTCTTAGAGCTCTAGAGCTCTAATATTGTTTACTCTCTTTGGACCTGTATCTTTAA 1409
Db 383 CTTCTTAGAGCTCTAGAGCTCTAATATTGTTTACTCTCTTTGGACCTGTATCTTTAA 324
Qy 1410 CTTCTTGTGTTAGTGTCTCTTTCCAGAAATGGAAGCTGTAAAGCTACAGATGCTTTACA 1469
Db 323 CTTCTTGTGTTAGTGTCTCTTTCCAGAAATGGAAGCTGTAAAGCTACAGATGCTTTACA 264
Qy 1470 AATGGAACCCCA 1481
Db 263 AATGGAACCCCA 252

RESULT 28
ACN44334/c
ID ACN44334 standard; DNA; 161334 BP.
XX
AC ACN44334;

XX 18-NOV-2004 (first entry)
DT
XX Human genomic sequence HCG32959.
DE
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
KW
XX Homo sapiens.
OS
XX WO2003073826-A2.
PN
XX 12-SEP-2003.
PD
XX 28-FEB-2003; 2003WO-US006235.
PF
XX 01-MAR-2002; 2002US-00087192.
PR
XX (SAGR-) SAGRES DISCOVERY.
PA
XX Morris DW;
PI
XX WPI; 2003-328604/31.
DR
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.
PT
XX Claim 1; SEQ ID NO 730; Opp; English.
PS
XX The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: this patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published
CC
SQ Sequence 161334 BP; 46474 A; 28689 C; 31549 G; 54602 T; 0 U; 20 Other;
Query Match 86.3%; Score 1277.4; DB 11; Length 161334;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 81; Indels 33; Gaps 4;
Qy 1 ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTTCTTACCCTTTTCTCTCACT 60
Db 102037 ATGGCCCTCCCTTATCATATTTTCTTTTACTGTTCTTACCCTTTTCTCTCACT 101978
Qy 61 GCACCCCTCCATGCTGTGTACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 101977 GCACCCCTCCATGCTGTGTACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 101918
Qy 121 AGCGCGCTTCCGGAATATTGATGCCCATCATATAGAGTTTATTAAGGAACTCC 180
Db 101917 ATGCGGCTTCCGGAATATTGATGCCCATCAATAGAGTTTACCTAAGAACTCC 101858
Qy 181 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATACTCTGCCACTCTTTGCATG 240
Db 101857 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATACTCTGCCACTCTTTGCATG 101798
Qy 241 CATGCAATATCATATTATGGACAGGAAATGATTATCTTAGTTGCTCTGAGACTT 300
Db 101797 CATGCAATATCATATTATGGACAGGAAATGATTATCTTAGTTGCTCTGAGACTT 101738
Qy 301 GGAGCCACTGCTGTTGGACTTTACTTACCCCATACAGTATGCTGATGGGGTGGAAAT 360
Db 101737 GGAGCCACTGCTGTTGGACTTTACTTACCCCATACAGTATGCTGATGGGGTGGAAAT 101678

Db 1023 CAAGAGTTTCTATGAGAAATGAGCGTCCCGAAATATTTGATGCCCCATTGCTATAGGAGT 1082
Qy 163 TTATCTAAGGGAATCCACCTTCACTGCCCCACACCCATATATGCCCGCAACTGCTATTAAC 222
Db 1083 TTATCTAAGGGAACCCCACTTCACTGCCCCACACCCATATATGCCCACTGCTATTAAC 1142
Qy 223 TCTGCACCTTTTGCATGACGAATACTCATTTATGACAGGGAATGATTAATCCT 282
Db 1143 TCTGCACCTTTTGCATGACGAATACTCATTTATGACAGGGAATGATTAATCCT 1202
Qy 283 AGTTGTCTGGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCCATACCAAGTATG 342
Db 1203 AGTTGTCTGGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCCATACCAAGTATG 1249
Qy 343 TCTGATGGGGTGGAAATCAAGGTGAGCAAGAGAAAACAAGTAAAGGAAGCAATCTCC 402
Db 1250 TCTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACATAAAGGAAGTAAATCTCC 1309
Qy 403 CAACCTGACCCGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTA 462
Db 1310 CAACCTGACCTGGGTACATAGCACCCCTAGCCCTACAAAGGACTAGATCTCTCAAAACTA 1369
Qy 463 CATGAACCCCTCGTACCCATACTCGCCTGGTGAGCCTATTTAATACCAACCCCTCACTCGG 522
Db 1370 CATGAACCCCTCGATACCCATACTGGCCTGGTAAGCCTATTTAATACCAACCCCTGACTGG 1429
Qy 523 CTCATAGAGTCTCAGCCCAAAACCCCTACTAATCTGTTGGATGTGCTCCCTCGACCTTC 582
Db 1430 CTCATAGAGTCTCGGCCCAAAACCCCTACTAATCTGTTGGATGTGCTCCCTCGACCTTC 1489
Qy 583 AGGCCATATATTTCAATCCTGTCTGTAACATGGAACAACTTCAGCACAGAAATAAAC 642
Db 1490 AGGCCATATATTTCAATCCTGTACTACCTGAACAAATGGAAACAATTCAGCACAGAAATAAAC 1549
Qy 643 ACCACTTCGGTTTGTAGTAGCCTCTGTTTCAATCTGGAATACCCATACCTCAAAAC 702
Db 1550 ACCACTTCGTTTGTAGTAGCCTCTGTTTCAATCTGGAATACCCATACCTCAAAAC 1606
Qy 703 CTCACCTGTGTAAATTTAGCAATATATAGACAAACAGAGTCCCAATGCATCAGGTGG 762
Db 1607 CTCACCTGTGTAAATTTAGCAATATATAGACAAACAGAGTCCCAATGCATCAGGTGG 1666
Qy 763 GTAACACCTCCACACGATGTCCTGCTACCTCAGGAATATTTTGTCTGTGTACC 822
Db 1667 GTAACACCTCCACACGATGTCCTGCTACCTCAGGAATATTTTGTCTGTGTACC 1726
Qy 823 TCAGCCTATCATTTGTTGAATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTAGTG 882
Db 1727 TCAGCCTATCATTTGTTGAATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTAGTG 1786
Qy 883 CCCCCTATGACCATCTACACTGAACAAAGATTTATACAATCATGTCTAGTAAAGCCCCAC 942
Db 1787 GCCCCTATGCCATCTACACTGAACAAAGATTTATACAATCATGTCTAGTAAAGCCCCAC 1846
Qy 943 ACACAAAGAGTACCCATCTCTCTTTTGTATCAGAGCAGGAGTGTAGCAGACTAGT 1002
Db 1847 ACACAAAGAGTACCCATCTCTCTTTTGTATTTGGAGCAGGAGTGTAGCAGGAGTGT 1906
Qy 1003 ACTGGCATTTGGCAGTATCAAACTCTACTCAGTTCTACTACAACTATCTCAAGAAATA 1062
Db 1907 ACTGGCATTTGGCAGTATCAAACTCTACTCAGTTCTACTACAACTATCTCAAGAAATA 1966
Qy 1063 AATGGTGACATGAAACAGGTCACTGACTCCCTGGTCACTTGAAGATCAATCACTACTCC 1122
Db 1967 AATGGTGACATGAAATGGGTGGTGTATCAGAGCAGGAGTGTAGCAGACTAGT 2026
Qy 1123 CTAGCAGCAGTACTCTTCAAAATCGAAGCTTTAGACTTGTAAACCGCCAAAGAGGG 1182
Db 2027 CTAGCAGCAGTACTCTTCAAAATCGAAGCTTTAGACTTGTAAACCGCGGAAGCGGG 2086
Qy 1183 GGAACCTGTTTATTTTATAGAGAAAGACGCTGTTTATTTATGTTAATCAATCCAGAAATGTC 1242
Db 2087 GGAACCTTATTTTATAGAGAAATGCTGTTTATGTTTATGTTAATCAATCCGGAATCATC 2146

Qy 1243 ACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGGAGCTTCAAAAC 1302
Db 2147 ACCGAGAAAGTTAAAGAAATTCAGGTCGAATATAACGTAGAGCAAGAGAGCTGCAAAAC 2206
Qy 1303 ACCGAAACCTCGGGGCCCTCTCAGCAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACT 1362
Db 2207 ACTGGACCTCGGGGCCCTCTCAGCAATGGATGCCCTGGATTCTCCCTTTCTTAGGACT 2266
Qy 1363 CTAGCAGCTCTAATATTTGTTACTCTCTTTGGACCCCTGTATCTTTAACTCCTTCTTTAAG 1422
Db 2267 CTAGCAGCTCTAATATTTGTTACTCTCTCTTTGGACCCCTGTATCTTTAACTCCTTCTTTAAG 2326
Qy 1423 TTTGTCTCTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA 1481
Db 2327 TTTGTCTCTTCCAGAAATCGAAGCTGTAAAGCTACAAATCGTCTTCAATGGAGCCCCA 2385

RESULT 30
AA114608
ID AA114608 standard; DNA; 1894 BP.
XX
AC AA114608;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #4541 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.
XX
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PI WPI; 2001-488901/53.
XX
DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 4541; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging of
XX diseases of the cervix, notably cervical cancer. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 76.4%; Score 1131.6; DB 4; Length 1894;
Best Local Similarity 92.4%; Pred. No. 0;

Qy	1244	CTGAGAAAGCTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAAACA	1300
Db	1065	CCGAGAAAGTTAAAGAAATTCAGAGTCGAATATACGTAGAGCAAGGAGCTGCNAACA	1124
Qy	1304	CCGAAAGCTGGGGCCTCCTCAGCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTC	1363
Db	1125	CTGACCCCTGGGGCCTCCTCAGCAATGGATGCCCTGGATTCTCCCTTCTTAGGACCTC	1184
Qy	1364	TAGCAGCTCTAAATATTGTTACTCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTTAAGT	1423
Db	1185	TAGCAGCTATAAATATTGTTACTCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTTAAGT	1244
Qy	1424	TTGTCTCTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA	1481
Db	1245	TTGTCTCTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA	1302
RESULT 31			
ABA56337			
ID	ABA56337 standard; DNA; 1894 BP.		
XX	ABA56337;		
XX	01-FEB-2002 (first entry)		
XX	Human foetal liver single exon nucleic acid probe #4642.		
XX	Human, foetal liver; gene expression; single exon nucleic acid probe; ss.		
XX	Homo sapiens.		
XX	WO200157277-A2.		
XX	09-AUG-2001.		
XX	30-JAN-2001; 2001WO-US000669.		
XX	04-FEB-2000; 2000US-0180312P.		
XX	26-MAY-2000; 2000US-0207456P.		
XX	30-JUN-2000; 2000US-00608408.		
XX	03-AUG-2000; 2000US-00632366.		
XX	21-SEP-2000; 2000US-0234687P.		
XX	27-SEP-2000; 2000US-0236359P.		
XX	04-OCT-2000; 2000GB-00024263.		
XX	(MOLE-) MOLECULAR DYNAMICS INC.		
XX	Penn SG, Hanzel DK, Chen W, Rank DR;		
XX	WPI; 2001-483447/52.		
XX	Human genome-derived single exon nucleic acid probes useful for analyzing		
XX	gene expression in human fetal liver.		
XX	Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.		
XX	The invention relates to a single exon nucleic acid probe for measuring		
XX	human gene expression in a sample derived from human foetal liver. The		
XX	single exon nucleic acid probes may be used for predicting, measuring and		
XX	displaying gene expression in samples derived from human fetal liver. The		
XX	present sequence is a single exon nucleic acid probe of the invention.		
XX	Note: The sequence data for this patent did not form part of the printed		
XX	specification, but was obtained in electronic format directly from WIPO		
XX	at ftp.wipo.int/pub/published_pct_sequences		
XX	Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;		
XX	Best Match		
XX	Query Local Similarity 76.4%; Score 1131.6; DB 4; Length 1894;		
XX	Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2		
Qy	164	TATCTAAGGAAATCTCACTTCTGCTCCACCCATATGCCCCGCAACTGCTATAACT	223

Db 61 CTGCCACTCTTTGCATGTCATGCAAAATCACTCAATTAATTTGGACAGGAAAAAGATTAATCCCA 120
 Qy 284 GTTGTCTCTGAGGACCTTGGAGCCACTGTCTGTTGGACTTACTTCAACCATACAGTATGT 343
 Db 121 GTTGTCTCTGAGGACCTTGGG-----GGACTCACTTCACTCATACAGTATGT 167
 Qy 344 CTGATGGGGTGGAAATTCAGAGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCC 403
 Db 168 CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAAATCTCC 227
 Qy 404 AACTGACCCGGGACATAGACACCCCTAGCCCTACAAAGGACTAGTCTCTCAAACTAC 463
 Db 228 AACTGACCTGGGTACATAGACACCCCTGGCCCTACAAAGGACTAGATCTCTCAAACTAC 287
 Qy 464 ATGAACCCCTCCGTAACCATCTCGCTGTGAGCCCTATTAAATACACACCTCACTCGGC 523
 Db 288 ATGAACCCCTCCATACCATCTAGCTGGCTGTGAGCCCTATTAAATACACACCTCACTGGG 347
 Qy 524 TCCATGAGGTCTCAGCCCAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCATCTCA 583
 Db 348 TCCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCATCTTA 407
 Qy 584 GGCATACATTTCAATCCCTGTTCTGAAACAATGGAACAACCTTCAGACACAGAAATAACA 643
 Db 408 GGCATACATTTCAATCCCTATACCTGAACAATGGAACAACCTTCAGACACAGAAATAACA 467
 Qy 644 CCATCTCCGTTTAGTAGGACCTCTGTTTCCATCTGGAATAAACCATACCTCAAAAC 703
 Db 468 CCATCTCTGTTTAGTAGGTCCTC---TTTCCATCTGGAATAAACCATACCTCAAAAC 524
 Qy 704 TCACCTGTGAAAAATTTAGCAATACTATAGACACAACCCAGCTCCCAATGATCAGGTGG 763
 Db 525 TCACCTGTGAAAAATTTAGCAATACTATAGACACAACCCAGCTCCCAATGATCAGGTGG 584
 Qy 764 TAACACCTCCACACAGAAATAGTGTGCTCAACCTCAGGAATATTTTTGTCTGTGGTACCT 823
 Db 585 TAACCTCCACACAGAAATAGTGTGCTCAACCTCAGGAATATTTTTGTCTGTGGTACCT 644
 Qy 824 CAGCTATCATTTGTTGAATGGCTCTTCAGAAATCTATGCTCTCTCTCTCTCTCTCTCT 883
 Db 645 CAGCTATCATTTGTTGAATGGCTCTTCAGAAATCTATGCTCTCTCTCTCTCTCTCTCT 704
 Qy 884 CCCCTATGACCATCTACACTGAACAAGATTTATCAATCATGCTGCTACCTAAAGCCCA 943
 Db 705 CCCCTATGACCATCTACACTGAACAAGATTTATCAATCATGCTGCTACCTAAAGCCCA 764
 Qy 944 ACAAAGAGTACCAATCTTCTCTTTGTTATCAGACGAGAGTCTAGGAGACTAGGTA 1003
 Db 765 ACAAAGAGTACCAATCTTCTCTTTGTTATTTGGAGCAGGAGTCTAGGCGGAGTAGCTA 824
 Qy 1004 CTGCAATGGGAGTATCAACACCTCTACTAGTTCTACTCAAACTATCTCAAGAAATAA 1063
 Db 825 CTGCAATGGGAGTATCAACACCTCTACTAGTTCTACTCAAACTATCTCAAGAAATAA 884
 Qy 1064 ATGTGATCATGGAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1123
 Db 885 ATGTGATCATGGAATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
 Qy 1124 TAGCAGCAGTGTCTCTTCAAAATCGAAGAGCTTTAGACTGCTTAACCGCAAAAGAGGG 1183
 Db 945 TAGCAGCAGTGTCTCTTCAAAATCGAAGAGCTTTAGACTGCTTAACCGCGGAAAGCGGG 1004
 Qy 1184 GAACCTGTTATTTTGGAGAGAAACGCTGTTATTAATGTTAAATCAATCCAGAAATGCTA 1243
 Db 1005 GAACCTGTTATTTTGGAGAGAAATGCTGTTATTAATGTTAAATCAATCCGGAATCATCA 1064
 Qy 1244 CTGAGAAATTTAAAGAAATTCGAGATCGAATACAAATGAGAGCAGAGCTTCAAAACA 1303
 Db 1065 CCGAGAAATTTAAAGAAATTCAGGTGCAATATACGTAGAGCAAGAGAGCTGCAAAACA 1124
 Qy 1304 CCGAACGCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGACCTC 1363
 Db 1125 CTGGACCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGATTTCTCCCTTCTTAGACCTC 1184

Qy 1364 TAGCAGCTCTAATATTTACTCTCTTTGGACCCCTGATATCTTTAACTCTCTTTAAGT 1423
 Db 1185 TAGCAGCTATAATATTTACTCTCTTTGGACCCCTGATATCTTTAACTCTCTTTAAGT 1244
 Qy 1424 TTGTCTCTTCCAGATTTGAAGCTGTAAAGCTACAGATGCTTTACAAATGGAAACCCCA 1481
 Db 1245 TTGTCTCTTCCAGATTCGAAGCAGTAAACTCAAAATCGTTCTTCAATGGAGCCCA 1302

RESULT 33

AB45822 standard; DNA; 1894 BP.

AC ABA45822;

XX 01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #4517.

XX Human; microarray; single exon probe; gene expression; breast; disease;
 KW cancer; ss.

XX Homo sapiens.

OS WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000662.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234587P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes, useful
 for measuring gene expression in sample derived from human breast,
 comprises number of single exon nucleic acid probes.

XX Claim 1; SEQ ID NO 4517; 327pp + Sequence Listing; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting the
 CC probes with a collection of detectably labelled nucleic acids derived
 CC from mRNA of human breast, and then measuring the label bound to each
 CC probe of the microarray. The probes are useful for verifying the
 CC expression of regions of genomic DNA predicted to encode proteins. They
 CC are useful for gene discovery, and for determining predisposition and/or
 CC prognosing breast disease. Gene expression analysis is useful for
 CC assessing the toxicity of chemical agents on cells. The microarray of
 CC this invention presents a far greater diversity of probes for measuring
 CC gene expression, with far less bias than expressed sequence tag
 CC microarrays. The method is suitable for rapid production of functional
 CC information from genomic sequence. The present sequence is a single exon
 CC nucleic acid probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match

76.4%; Score 1131.6; DB 4; Length 1894;

Best Local Similarity 92.4%; Pred. No. 0;			
Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;			
QY	164	TATCTAAGGGAACCTCCACTTCACTGCCCCACCACTATGCCCCGCAACTGCTATAACT	223
Db	1	TATCTAAGGGAACCCCACTTCACTGCCACACCATATGCCCACTGCTATAACT	60
QY	224	CTGCCACTCTTTGTCATGCAATCACTCAATATTTGGACAGGGAATGATTATCTTA	283
Db	61	CTGCCACTCTTTGTCATGCAATCACTCAATATTTGGACAGGGAATGATTATCTTA	120
QY	284	GTGTCTCTGGAGACTTTGAGCCACTGTCTGTGTGACTTTACTTCCACCATACAGTATGT	343
Db	121	GTGTCTCTGGAGACTTGA-----GGACTCACTTCACTCATACAGTATGT	167
QY	344	CTGATGGGGTGAATTCAGGTTCAGGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCC	403
Db	168	CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACATAAAGGAAGTAAATCTCCC	227
QY	404	AACTGACCCGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTAC	463
Db	228	AACTGACCTGGGTACATAGCACCCCTGGCCCTTACAAAGGACTAGATCTCTCAAAACTAC	287
QY	464	ATGAAACCTTCGGTACCCCTACTCGCCTGGTGAGCCTATTATATACCACTCACTCGGC	523
Db	288	ATGAAACCTTCATATACCCCTACTGGCTGGTAAAGCCTATTATATACCACTCACTGGGC	347
QY	524	TCCATGAGTCTCAGCCCCAAACCTTAACTAATGTTGGATGTCCTCCCTCGCACTTCA	583
Db	348	TCCATGAGTCTCGGCCCAAAACCTTAACTAATGTTGGATGTCCTCCCTCGCACTTCA	407
QY	584	GGCCATACATTTCAATCCCTGTCTCGACAACTGGAACAACTTCAGCACAGAAATAACA	643
Db	408	GGCCATACATTTCAATCCCTGTCTCGACAACTGGAACAACTTCAGCACAGAAATAACA	467
QY	644	CAACTTCCGTTTGTAGGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAAAC	703
Db	468	CAACTTCTGTTTGTAGTGTCTC---TTTCCAACTGGGAATAACCCATACCTCAAAAC	524
QY	704	TCACTGTGTAAATTTAGCAATATATATAGACACAAACAGCTCCCAATCATCAGTGGG	763
Db	525	TCACTGTGTAAATTTAGCAATATATATAGACACAAACAGCTCCCAATCATCAGTGGG	584
QY	764	TACACCTCCCAACCAATGCTGCTACCTCCCTCAGGAATATTTTTGTCTGTGTACCT	823
Db	585	TAACTCTCCCAACCAATGCTGCTACCTCCCTCAGGAATATTTTTGTCTGTGTACCT	644
QY	824	CAGCCTATCATTTGTTGAATGGCTCTTCAGAACTCTATGCTCTCCTCTCATTTTAGTGC	883
Db	645	CAGCCTATCATTTGTTGAATGGCTCTTCAGAACTCTGCTGCTCTCATTTTAGTGG	704
QY	884	CCCTATGACATCTACACTGAACAAGATTTATCAATCATGTCGTACCTAAGCCCCACA	943
Db	705	CCCTATGCCCCATCTACACTGAACAAGATTTATCAATCATGTCATACCTAAGCCCCCA	764
QY	944	ACAAAGAGTACCATTTCTCTTTTGTATCAGACGAGGTCGTAGCAGACTAGGTA	1003
Db	765	ACAAAGAGTACCATTTCTCTTTTGTATTTGGACGAGAGTGTAGGGGAGTAGCTA	824
QY	1004	CTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAA	1063
Db	825	CTGGCATTTGGCGTATCACAACTCTACTCAGTTCTTACTACAACTGTCTCAAGAAATAA	884
QY	1064	ATGGTACATGGAACAGGTCACTGACTCCCTGTGCTACCTGCGAGATCACTTAATCTCC	1123
Db	885	ATGGTACATGGAATGGGTGCTGTATACCTGCTGCTACCTTGCAGATCACTTAATCTCC	944
QY	1124	TAGCAGCAGTGTCTTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAGAGGGG	1183
Db	945	TAGCAGCAGTGTCTTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGGGAAGCGGG	1004
QY	1184	GAACCTGTTTATTTTAGGAGAGAAACGCTGTTTATTTATGTTAATCAATCCAGAAATGTCA	1243

Db	1005	GAACCTTTTATTTTATAGAGGAAAAATGCTGTTGTTATGTTTAAATCAATCCGGAATCATCA	1064
QY	1244	CTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACA	1303
Db	1065	CCGAGAAAGTTAAAGAAATTCAGGTGCAATATATACGTAGAGCAAGAGGCTGCANAACA	1124
QY	1304	CCGAAACGCTGGGGCTCTCTCAGCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTC	1363
Db	1125	CTGGACCTCGGGCTCTCTCAGCCNAATGGATGCCCTGGATTCTCCCTTCTTAGGACCTC	1184
QY	1364	TAGCAGCTCTAATATATGTTACTCTCTTTGGACCTGTATCTTTAACCTCTTGTAAAGT	1423
Db	1185	TAGCAGCTCTAATATATGTTACTCTCTTTGGACCTGTATCTTTAACCTCTTGTAAAGT	1244
QY	1424	TTGTCTCTCCAGAAATGGAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAACCCCA	1481
Db	1245	TTGTCTTTTTCAGAAATCGAAGCAGTAAACCTACAAATCGTTCTTCAATTTGAGCCCCA	1302
RESULT 34			
ABA25978			
ID	ABA25978	standard; DNA; 1894 BP.	
XX	ABA25978;		
AC	AC		
XX	XX		
DT	DT		
XX	23-JAN-2002 (first entry)		
DE	DE	Probe #4444 for gene expression analysis in human heart cell sample.	
XX	XX		
KW	KW	Human; gene expression; heart; microarray; vascular system; probe;	
KW	KW	cardiovascular disease; hypertension; cardiac arrhythmia;	
XX	XX	congenital heart disease; ss.	
OS	OS	Homo sapiens.	
XX	XX		
PN	PN	WO200157274-A2.	
XX	XX		
PD	PD	09-AUG-2001.	
XX	XX		
PF	PF	30-JAN-2001; 2001WO-US000666.	
XX	XX		
PR	PR	04-FEB-2000; 2000US-0180312P.	
PR	PR	26-MAY-2000; 2000US-0207456P.	
PR	PR	30-JUN-2000; 2000US-00608408.	
PR	PR	03-AUG-2000; 2000US-00632366.	
PR	PR	21-SEP-2000; 2000US-0234687P.	
PR	PR	27-SEP-2000; 2000US-0236359P.	
PR	PR	04-OCT-2000; 2000GB-00024263.	
XX	XX	(MOLE-) MOLECULAR DYNAMICS INC.	
XX	XX		
PI	PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX	XX		
DR	DR	WPI; 2001-48899/53.	
XX	XX		
PT	PT	Single exon nucleic acid probes for analyzing gene expression in human hearts.	
XX	XX		
PS	PS	Claim 1; SEQ ID NO 4444; 530pp; English.	
XX	XX		
CC	CC	The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX	XX	Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;	

Qy	1184	GAACCTGTTTATTTTAGGAGAGAGACGCTGTTATTATGTTAATCAATCAGAAATTCGTCA	124
Db	1005	GAACCTTTTTTATTTTAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA	1064
Qy	1244	CTGAGAAAGTTAAAGAAAATTCGAGATCGAATCAAAATGTAGAGCAGAGGAGCTTCAAAAACA	1303
Db	1065	CCGAGAAAGTTAAAGAAATTCAGGTCGAATATAACGTAGACGAAAGGAGCTGCAAAAACA	1124
Qy	1304	CCGAAACGCTGGGGCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCTCTTCTTAGGACCTC	1363
Db	1125	CTGGACCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCTTCTTAGGACCTC	1184
Qy	1364	TAGCAGCTCTAAATATTTGTTTACTCCTCTTTGGAGCCCTGTATCTTTAACTCCTCTGTTAAGT	1423
Db	1185	TAGCAGCTATAATATTTGTTTACTCCTCTTTGGAGCCCTGTATCTTTAACTCCTCTGTTAAGT	1244
Qy	1424	TTGTCCTCTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA	1481
Db	1245	TTGTCCTCTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA	1302
RESULT 35			
AAK30018			
ID	AAK30018 standard; DNA; 1894 BP.		
XX	AAK30018;		
XX	06-NOV-2001 (first entry)		
DT	Human bone marrow expressed single exon probe SEQ ID NO: 4575.		
XX	Human; bone marrow expressed exon; gene expression analysis; probe;		
KW	microarray; cancer; leukaemia; lymphoma; myeloma; ss.		
KW	Homo sapiens.		
OS	WO200157276-A2.		
XX	09-AUG-2001.		
PN	30-JAN-2001; 2001WO-US000668.		
XX	04-FEB-2000; 2000US-0180312P.		
PF	26-MAY-2000; 2000US-0207456P.		
XX	30-JUN-2000; 2000US-00608408.		
XX	03-AUG-2000; 2000US-00632366.		
XX	21-SEP-2000; 2000US-0234687P.		
XX	27-SEP-2000; 2000US-0236359P.		
XX	04-OCT-2000; 2000GB-00024263.		
PA	(MOLE-) MOLECULAR DYNAMICS INC.		
XX	Penn SG, Hanzel DK, Chen W, Rank DR;		
FI	WPI; 2001-488900/53.		
XX	Human genome-derived single exon nucleic acid probes useful for analyzing		
XX	gene expression in human bone marrow.		
PT	Example 4; SEQ ID NO 4575; 659pp + Sequence Listing; English.		
PS	The present invention provides a number of single exon nucleic acid		
XX	probes which are derived from genomic sequences expressed in the human		
CC	bone marrow. They can be used to measure gene expression in bone marrow		
CC	samples, which may enable the improved diagnosis and treatment of cancers		
CC	such as lymphoma, leukaemia and myeloma. The present sequence is one of		
CC	the probes of the invention		
XX	Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;		
XX	Query Match 76.4%; Score 1131.6; DB 4; Length 1894;		
XX	Best Local Similarity 92.4%; Pred. No. 0;		
XX	Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps		

QY 164 TATCTAAGGGAACCTCCACCTTCACTGCCACACCATATGCCCCGCAACTGCTATACT 223
DB 1 TATCTAAGGGAACCTCCACCTTCACTGCCACACCATATGCCCCCAAACTGCTATACT 60
QY 224 CTGCCACTCTTTGCGATGCAATCACTCACTATTGGAAGGGAATGATTATCTTA 283
DB 61 CTGCCACTCTTTGCGATGCAATCACTCACTATTGGAAGGGAATGATTATCTTA 120
QY 284 GTTGCTCTGGAGCACTTGAGGCACTGTCTGTGGAAGGCACTTACCCATACAGTATGT 343
DB 121 GTTGCTCTGGAGCACTTGGA-----GGACTCACTTCACTCATACAGTATGT 167
QY 344 CTGATGGGGTGGAACTTCAAGGTCAGGCAAGAGAAACAACTGGAAGGCAAACTCC 403
DB 168 CTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACAACTGGAAGGCAAACTCC 227
QY 404 AACTGACCCGGGACATAGCACCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTAC 463
DB 228 AACTGACCTGGGTACATAGCACCCTGGCCCTACAAAGGACTAGATCTCTCAAACTAC 287
QY 464 ATGAACCTCGTACCCATCTCGCTGGTGAAGCTATTTAATACCACTCCTCCTGGC 523
DB 288 ATGAACCTCGTACCCATCTCGCTGGTGAAGCTATTTAATACCACTCCTCCTGGC 347
QY 524 TCCATGAGCTCAGCCCAAAACCTTACTAAGTGTGGATGCTCCCTCCCTGCACTTCA 583
DB 348 TCCATGAGCTCAGCCCAAAACCTTACTAAGTGTGGATGCTCCCTCCCTGCACTTCA 407
QY 584 GGCCATACATTTCAATCCCTGTTCTCTGAACTGGAACAACTTACGACAGAAATAACA 643
DB 408 GGCCATACATTTCAATCCCTTATACCTGAACAACTTACGACAGAAATAACA 467
QY 644 CCATCTCGTTTGTAGGACCTCTGTTTCCATCTGGAATTAACCCATACCTCAAACT 703
DB 468 CCATCTCTGTTTGTAGGACCTCTGTTTCCATCTGGAATTAACCCATACCTCAAACT 524
QY 704 TCACCTGTGTAAATTTAGCAATCTATAGACACCACTGCTCCCAATCATCAGTGGG 763
DB 525 TCACCTGTGTAAATTTAGCAATCTATAGACACCACTGCTCCCAATCATCAGTGGG 584
QY 764 TAACACCTCCCAACAGATGCTGCTACCTCAGGAATAATTTTGTGTGTGCTACCT 823
DB 585 TAACCTCCCAACAGATGCTGCTACCTCAGGAATAATTTTGTGTGTGCTACCT 644
QY 824 CAGCCTATCATGTTTGAATGGCTCTTCAAGATCTATGCTTCTCTCATCTTCTAGTGC 883
DB 645 CAGCCTATCATGTTTGAATGGCTCTTCAAGATCTATGCTTCTCTCATCTTCTAGTGC 704
QY 884 CCCCTATGACCTCTACACTGAACAAGATTTATACAACTATGCTGCTTACGCCCA 943
DB 705 CCCCTATGACCTCTACACTGAACAAGATTTATACAACTATGCTGCTTACGCCCA 764
QY 944 ACAAAGAGTACCCATCTTCTCTTTGTTATCAGACGAGGAGTGTAGGAGACTAGTGA 1003
DB 765 ACAAAGAGTACCCATCTTCTCTTTGTTATGAGCAGGAGTGTAGGAGACTAGTGA 824
QY 1004 CTGGCATTTGGCAGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATA 1063
DB 825 CTGGCATTTGGCAGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATA 884
QY 1064 ATGGTCACATGGAACAGGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
DB 885 ATGGTCACATGGAACAGGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
QY 1124 TAGCAGCAGTGTCTTCAAAATCGAAGACTTTAGACTGTGCTTAAACCGCAAGAGGG 1183
DB 945 TAGCAGCAGTGTCTTCAAAATCGAAGACTTTAGACTGTGCTTAAACCGCAAGAGGG 1004
QY 1184 GAACCTGTTTATTTTAGGAGAGAGCTGTTTATTTATGTTATCAATCCAGATTTGCTCA 1243
DB 1005 GAACCTGTTTATTTTAGGAGAGAGAGCTGTTTATTTATGTTATCAATCCGGAATCATCA 1064

QY 1244 CTGAGAAAGTTTAAAGAAATTCAGATCGAATACAAATGTAGAGCAGAGGCTTCAAACA 1303
DB 1065 CCGAGAAAGTTTAAAGAAATTCAGATCGAATACAAATGTAGAGCAGAGGCTTCAAACA 1124
QY 1304 CCGAACGCTGGGGCTCTCTAGCCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTC 1363
DB 1125 CTGACCCCTGGGGCTCTCTAGCCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTC 1184
QY 1364 TAGCAGCTCTAATATTGTTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTAAAGT 1423
DB 1185 TAGCAGCTCTAATATTGTTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTAAAGT 1244
QY 1424 TTGTCTCTTCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
DB 1245 TTGTCTCTTCAGAAATGAAGCTGTAAAGCTACAAATGCTTCTTCAAAATGGAGCCCA 1302
RESULT 36
AAK04516
ID AAK04516 standard; DNA; 1894 BP.
XX AAK04516;
AC AAK04516;
XX 05-NOV-2001 (first entry)
DT 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe SEQ ID NO: 4507.
DE Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX Homo sapiens.
OS Homo sapiens.
XX WO200157275-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000667.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human brains.
PT Example 4; SEQ ID NO 4507; 650pp + Sequence Listing; English.
PS The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match 76.4%; Score 1131.6; DB 4; Length 1894;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;
QY 164 TATCTAAGGGAACCTCCACCTTCACTGCCACACCATATGCCCCGCAACTGCTATACT 223

1	TATCTAAGGGNACCCCCACCTTCACTGCCACACCCATATGCCCCACAATCGCTAACT	60
224	CTGCCACTCTTTGCATGCATGCAAACTCATTTATTTGGACAGGGAATAATGATTAATCTCTA	283
61	CTGCCACTCTTTGCATGCATGCAAACTCATTTATTTGGACAGGGAATAATGATTAATCTCTA	120
284	GTGTGCTCTGGAGACTTTGGAGCCACTGTCTGTGTGGACTTACTTTCAACCCATACAGTATGT	343
121	GTGTGCTCTGGAGACTTTGGA-----GGACTCACTTCACTACACAGTATGT	167
344	CTGATGGGGTGGAATTCAGAGTTCAGCAAGAGAAAAAAGTAAAGGAAGCAATCTCCC	403
168	CTGATGGGGTGAGATTCAAGATCAGCAACAGAAAAACAATAAAGGAAGTAAATCTCCC	227
404	AACGTACCCGGGACATAGCACCCCTAGCCCTTACAAAGGACTAGTCTCTCAAAACTAC	463
228	AACGTACCTGGGTACATAGCACCCCTGGCCCTTACAAAGGACTAGTCTCTCAAAACTAC	287
464	ATGAAACCCCTCCGTACCCATACTCGCTGTGTAGCCCTATTAAATACACCTCACTGGC	523
288	ATGAAACCCCTCCATACCCATACTGGCTGTGTAGCCCTATTAAATACACCTGACCTGGC	347
524	TCCATGAGGTCTCAGCCCAAAAACCTACTAACTGTGTGGATGTGGCTCCCTCGCACTTCA	583
348	TCCATGAGGTCTCGGCCAAAAACCTACTAACTGTGTGGATGTGGCTCCCTCGCACTTCA	407
584	GGCCATACATTTCAATCCCTGTTCCGTAAACAATGGAACAACCTTCAGCACAGAAATAACA	643
408	GGCCATACATTTCAATCCCTTATACCTGAAACAATGGAACAACCTTCAGCACAGAAATAACA	467
644	CCACTTCGGTTTGTAGTAGGACCTCTTGTCTTCCAACTCTGGAATAAACCCATCAAC	703
468	CCACTTCGGTTTGTAGTAGGCTCTC-----TTTCCAACTCTGGAATAAACCCATCAAC	524
704	TCACCTGTGTAAAAATTTAGCAATACTATAGACACAACACAGGTCCCAATGATCAGGTGG	763
525	TCACCTGTGTAAAAATTTAGCAATACTATAGACACAACACAGGTCCCAATGATCAGGTGG	584
764	TAAACCTCCACACAGTAAGTCTGCTACCTCACCAAGTAATTTTTTGTCTGTGGTACCT	823
585	TAACTCTCCACACAGTAAGTCTGCTACCTCACCAAGTAATTTTTTGTCTGTGGTACCT	644
824	CAGCTATCATTTGTGAATGGCTCTTCAGAACTATGTGTCTCTCTCAATCTTTAGTGC	883
645	CAGCTATCATTTGTGAATGGCTCTTCAGAACTATGTGTCTCTCTCAATCTTTAGTGC	704
884	CCCCTATGACCATCTACACTGAAACAAGATTTATCAATCATGTGCTACTTAAGCCCCACA	943
705	CCCCTATGACCATCTACACTGAAACAAGATTTATCAATCATGTGCTACTTAAGCCCCACA	764
944	ACAAAAGAGTACCATCTTCTCTTTGTATCAGACGAGGAGTCTAGGACAGCTAGCTA	1003
765	ACAAAAGAGTACCATCTTCTCTTTGTATTTGGACGAGGAGTCTAGGCGGAGTAGCTA	824
1004	CTGGCATTTGGCAGTATCAACACTCTACTCAGTTTCTACTACAACACTATCTCAAGAAATAA	1063
825	CTGGCATTTGGCGGTATCACACCTCTACTCAGTTTCTACTACAACACTATCTCAAGAAATAA	884
1064	ATGTGACATGGAAACAGGTCACTGACTCCCTGGTCACTTGTGCAAGATCAACTTAACCTCC	1123
885	ATGTGTACATGGAAATGGGTGCTGATACCTGGTCACCTTGTGCAAGATCAACTTAACCTCC	944
1124	TAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAAGAGGGG	1183
945	TAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAAGAGGGG	1004
1184	GAACTGTTTATTTTGTAGGAGAAGAACGCTGTATTATTTATTTAAATCAATCCAGAAATGTC	1243
1005	GAACTGTTTATTTTGTAGGAGAAGAACGCTGTATTATTTATTTAAATCAATCCAGAAATGTC	1064
1244	CTGAGAAAAGTTAAAGAAATTCGAGATCGAAATACAATGTAGAGCAGAGGAGCTTCAAAACA	1303

Db	1065	CCGAGAAAGTTAAAGAAATTCAGAGTCGCGAATAAATCGTAGCAGCAAGAGGAGCTGCAAAACA	1122
Qy	1304	CCGAAAGCGCTGGGGCCTCTCAGCCAATGGATGCCCTGGGTCTCCCTCTTTAGGACCTC	1363
Db	1125	CTGGACCTTGGGGCCTCTCAGCCAATGGATGCCCTGGATTCTTCCCTCTTTAGGACCTC	1184
Qy	1364	TAGCAGCTCTAAATATTGTTACTCCTCTTTGGACCTGTATCTTTAACTCCTTGTTAAGT	1423
Db	1185	TAGCAGCTAATAATTGTTACTCCTCTTTGGACCTGTATCTTTAACTCCTTGTTAAGT	1244
Qy	1424	TTGTCTCTTCCAGAAATTGAAGCTGTAAGCTACAGATGGTCTTACAATATGGAAACCCCA	1481
Db	1245	TTGTCTTTTCCAGAAATCGAAGCAGTAAACTACAAATCGTTCTTTCAATATGAGGCCCA	1302

DESCRIPT 37

```

RES001 3,
ABS29670
ID ID ABS29670 standard; DNA; 1894 BP.
XX
XX AC ABS29670;
XX
XX DT 25-FEB-2003 (first entry)
XX
XX DE Human liver single exon probe, SEQ ID No 4660.
XX
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; SS.
XX
XX OS Homo sapiens.
XX
XX PN WO200157273-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US0000664.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00832366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI; 2001-488898/53.

```

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.

Claim 1. SEQ ID NO 4660: 658pp: English.

The invention relates to a single exon nucleic acid probe (GENP) (I) for measuring human gene expression in a sample derived from human adult liver, comprising one of 13109 defined nucleotide sequences given in the specification (or complements/ fragments). The probe hybridises at high stringency to a nucleic acid molecule expressed in the human adult liver. (I) may be used for predicting, measuring and displaying gene expression in samples derived from human adult liver. The genes identified may be involved in genetic liver diseases such as cirrhosis, hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is associated with coronary heart disease. ABS25011-ABS51005 represent human liver single exon nucleic acid probes of the invention. Note: The sequence information for this patent does not appear in the printed specification but was obtained in electronic format directly from WIPO at <http://www.int/pub/patent/patseq>

Sequence 1894 BP: 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match	76.4%;	Score 1131.6;	DB 4;	Length 1894;
Best Local Similarity	92.4%;	Prod. No. 0;		
Matches 1218;	Conservative	0;	Mismatches 84;	Indels 16; Gaps 2;
Qy	164	TATCTAAGGGAAATCCACCTTCTACTGCCACACCCATATGCCCGCAACTGCTATAACT	223	
Db	1	TATCTAAGGGAAACCCCACTTCTACTGCCCAACCCATATGCCCAACAACTGCTATAACT	60	
Qy	224	CTGCCACTTTTGGCATGCATAAATCTCATTTATGGACAGGGAAATGATTAATCCTA	283	
Db	61	CTGCCACTTTTGCATGCATAAATCTCATTTATGGACAGGAAACGATTAATCCCA	120	
Qy	284	GTTCCTCTGGAGGACTTGGGCCACTGCTGTTGGACTTACTTCCCATACCAAGTAGTGT	343	
Db	121	GTTCCTCTGGAGGACTTGGGA-----GGACTCACTTCATCATACCAAGTAGTGT	167	
Qy	344	CTGATGGGGTGGAAATTCAGGTCAGGCAAGAGAAAAACAAAGTAAAGGAACAATCTCCC	403	
Db	168	CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACAATAAAGGAAGTAATCTCCC	222	
Qy	404	AAC TGACCCGGGACATAGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAACTAC	463	
Db	228	AAC TGACCTGGGTACATAGCACCCCTGGCCCTCAAAAGGACTAGATCTCTCAAAACTAC	287	
Qy	464	ATGAAACCTCCGTACCCATACTCGCTGTGGAGCCTATTTAATACACACCTCTACTCGGC	523	
Db	288	ATGAAACCTCCATACCCATACTGGCTGGTAAGCCTATTTAATACACACCTGACTGGC	347	
Qy	524	TCCATAGGCTTCAGGCCAAACCCCTACTAACTGTTGATGTGCTGCCCTGCACTTCA	583	
Db	348	TCCATAGGCTTCGGGCCAAACCCCTACTAACTGTTGATGTGCTGCCCTGCACTTTA	407	
Qy	584	GGCATACATTTCAATCCCTGTCCTGAAACAATGGAACAACCTTCAGACACAGAAATAACA	643	
Db	408	GGCCTACATTTCAATCCCTATACCTGAACAANTGGAACAACCTTCAGACACAGAAATAACA	467	
Qy	644	CCACTTCGTTTATAGTAGGACCTCTGTGTTTCCAACTCTGGAATAACCCATACCTCAAAAC	703	
Db	468	CCACTTCGTTTATAGTAGGTCCTC---TTTCCAACTCTGGAATAACCCATACCTCAAAAC	524	
Qy	704	TCACCTGTGTAATAATTTAGCAATACTATAGCACAAACAGCTGCCAATGCATCAGGTGG	763	
Db	525	TCACCTGTGTAATAATTTAGCAATACTATAGCACAGCCAACTCCCAATGCATCAGGTGG	584	
Qy	764	TAAACCTCCACACAGAAATAGTCGCTTACCTCAGGAATATTTTTGTCTGTGGTACCT	823	
Db	585	TAACCTCTCCACACGAATAGTCGCTTACCTCAGGAATATTTTTGTCTGTGGTACCT	644	
Qy	824	CAGCTATCATTTTGAATGGCTCTTCAGAAATCATGTGCTTCTCTCATTTCTTAGTGC	883	
Db	645	CAGCTATCATTTTGAATGGCTCTTCAGAAATCTGTGTGCTTCTCTCATTTCTTAGTGG	704	
Qy	884	CCCCTATGACCATCTACACTGAAACAAGATTTATACAATCATGTGCTACCTTAAGCCCCACA	943	
Db	705	CCCCTATGCCCAATCTACACTGAAACAAGATTTATACAATCATGTGCTACCTTAAGCCCCGCA	764	
Qy	944	ACAAAAGATACCCATTTCTTCCTTTTGTATTCAGACAGGAGTGCTAGGCAGACTAGGTA	1003	
Db	765	ACAAAAGATACCCATTTCTTCCTTTTGTATTTGGAGCAGGAGTGCTAGGCAGGAGTAGCTA	824	
Qy	1004	CTGCAATTGGCAGTATCACACCTCTACTCAGTTTCTACTACAACACTATCTCAAGAATAA	1063	
Db	825	CTGCAATTGGCGGTATCACAACTCTACTCAGTTTCTACTACAACACTGTCTCAAGACTAA	884	
Qy	1064	ATGGTGATCTGGAAACAGGTCACTGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCC	1123	
Db	885	ATGGTGATCTGGAAATGGGTGCTGATACCTGGTCACTTTGCAAGATCAACTTAACCTCC	944	
Qy	1124	TAGCAGAGTAGTCCCTTCAAAATCGAAGCTTTAGACTTGTCTTAACCGCCAAAGAGGG	1183	
Db	945	TAGCAGAGTAGTCCCTTCAAAATCGAAGCTTTAGACTTGTCTTAACCGCCGGAAGCGGG	1004	
Qy	1184	GAACCTGTTTATTTTTAGGAGAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATGTCA	1243	

SQ Sequence 1894 Bp; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 76.4%; Score 1131.6; DB 5; Length 1894;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;

164 TATCTAAGGAAACTCCACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAACT 223
1 TATCTAAGGAAACCCCACTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAACT 60
224 CTGCCACTCTTTCATGATGCAAAATCACTATTATGGACAGGAAATTAATTCCTTA 283
61 CTGCCACTCTTTCATGATGCAAAATCACTATTATGGACAGGAAATTAATTCCTTA 120
284 GTTGTCTCTGAGGACCTGAGCCACTGTCTGTTGGACTTACTTCAACCATACCAATATGT 343
121 GTTGTCTCTGAGGACCTGGA-----GGACTCACTTCACTATACCAATATGT 167
344 CTGATGGGGGTGGAATTCAGAGGTGAGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCC 403
168 CTGATGGGGGTGAGTTCAGATCAGGCACAGAAACACATATAAGGAAGTAACTCTCCC 227
404 AACTGACCCGGGACATAGCACCCCTAGCCCTTACAAAGCACTAGTCTCTCAAACTAC 463
228 AACTGACCTGGGTACATAGCACCCCTGGCCCTTACAAAGCACTAGATCTCTCAAACTAC 287
464 ATGAACCCCTCGTACCCTACTCGCTGGTGGAGCCCTATTATATACCAACCTCACTCGGC 523
288 ATGAACCCCTCCATACCCTACTGGCTGGTGGAGCCCTATTATATACCAACCTCACTGGGC 347
524 TCCATGAGGTCTCAGCCCAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCA 583
348 TCCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCA 407
584 GGCCATACATTTCAATCCCTGTTCTCGAACAAATGGAACAATTCAGACAGAAATAAACA 643
408 GGCCATACATTTCAATCCCTATACCTGAACAATGGAACAATTCAGACAGAAATAAACA 467
644 CCACCTCGTTTGTAGTAGGACCTCTGTTCCCAATCTGGAATTAACCACTACCTCAAAAC 703
468 CCACCTCTGTTTGTAGTAGGCTCTC---TTTCCCAATCTGGAATTAACCACTACCTCAAAAC 524
704 TCACCTGTGTAATTTAGCAATATAGACACAACAGCTCCCAATGCATCAGGTGGG 763
525 TCACCTGTGTAATTTAGCAATATAGACACAACAGCTCCCAATGCATCAGGTGGG 584
764 TAACACCTCCACACGAAATAGTCTGCTCAACCTCAGGAATATTTTGTCTGTGTACCT 823
585 TAACCTCTCCACACGAAATAGTCTGCTCAACCTCAGGAATATTTTGTCTGTGTACCT 644
824 CAGCTATCATTTGTAATGGCTCTTCAGNAATCATGTGCTTCTCTCATTTCTTAGTGC 883
645 CAGCTATCATTTGTAATGGCTCTTCAGAAATCATGTGCTTCTCTCATTTCTTAGTGC 704
884 CCCCTATGACCATCTACACTGAACAAGATTTATACATCATGTCTACCTTAAGCCCAACA 943
705 CCCCTATGCCCATCTACACTGAACAAGATTTATACATCATGTCTACCTTAAGCCCAACA 764
944 ACAAAGAGTACCCATCTCTCTTTGTTATCAGACGAGGAGTCTAGGACAGTAGGTA 1003
765 ACAAAGAGTACCCATCTCTCTTTGTTATTTGGAGCAGGAGTCTAGGCGGAGTAGCTA 824
1004 CTGGCATGGCAGTATCAACCTCTACTAGTTCTACTACAACTATCTCAGAAATATAA 1063
825 CTGGCATGGCAGTATCAACCTCTACTAGTTCTACTACAACTGTCTCAAGAACTAA 884
1064 ATGGTGACATGGAACAGTCTGACTCCCTGGTCACCTTGCAGAGATCAACTTAACCTCCC 1123
885 ATGGTGACATGGAATGGGTCTGATACCTTGGTCACCTTGCAGAGATCAACTTAACCTCCC 944
1124 TAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAAAGAGGG 1183
945 TAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCGGAAGCGGG 1004

1184 GAACCTGTTTATTTTAGGAGAAACGCTGTTATATGTTATATGTTATCAATCCAGATGTCA 1243
1005 GAACCTGTTTATTTTAGGAGAAACGCTGTTATGTTATGTTATCAATCCGGAATCATCA 1064
1244 CTGAGAAAGTTAAAGAAATTCAGATCGAATCAATGTTAGACAGAGGAGCTTCAAAACA 1303
1065 CCGAGAAAGTTAAAGAAATTCAGAGTCAATATAACGTAGAGCAAGGAGCTGCAAAACA 1124
1304 CCGAAGCGTGGGCTCTCCAGCAATGATGATGCGCTGGGTTCTCCCTCTTTAGGACCTC 1363
1125 CTGACCCCTGGGCTCTCTCAGCCATGATGATGCGCTGGATCTCCCTCTTTAGGACCTC 1184
1364 TAGCAGCTCAATATGTTTACTCTCTTTGGACCCCTGATCTTTAACTCTCTTTAAAGT 1423
1185 TAGCAGCTAATATGTTTACTCTCTTTGGACCCCTGATCTTTAACTCTCTTTAAAGT 1244
1424 TTGCTCTTCCAGATTTGAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAAACCCCA 1481
1245 TTGCTCTTCCAGATTCGAAGCAGTAAACTACAAATCGTCTTCAATGGAGGCCCA 1302

RESULT 39

ABS04589
ID ABS04589 standard; DNA; 1894 Bp.

XX ABS04589;

XX 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe from lung SEQ ID No 4580.

XX Human; da; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX Claim 1; SEQ ID NO 4580; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their

complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes ; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemangioendothelioma, histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences

Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match	Score	DB	Length
Best Local Similarity	76.4%	131.6	1894
Matches 1218; Conservative	92.4%	Pred. NO. 0;	
Matches 1218; Conservative	0;	Mismatches 84;	Indels 16; Gaps 2

164	Qy	TATCTAAGGGAAA	CTCCACCTTCACTGCCCCACACCCATATGCCCCGGCAACTGCTATAACT	223
1	Db	TATCTAAGGGAAACCCCACTTCACTGCCCCACCCATATGCCCCCAACCTGCTATAACT		60
224	Qy	CTGCCACTCTTTGGATGCATGCAAAATACTCATTTATTGGACAGGGAAAAATGATTAACTCTTA	283	
61	Db	CTGCCACTCTTTGCATGTATGCAAAATACTCATTTATTGGACAGGAAAAACGATTAATCCCA		120
284	Qy	GTTGTCTCTGGAGGACTTGGAGCCCACTGTCTGTGTGGACTTACTTCCACCATACCAGTAGT	343	
121	Db	GTTGTCTCTGGAGGACTTGGAA-----GGACTCACTTCACTCATACCAGTAGT		167
344	Qy	CTGATGGGGTGGAAATTCAAAGTTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCC	403	
168	Db	CTGATGGGGTGGAGTTTCAAAGTATCAGGCACAGAAAAACACATAAAGGAAGTAATCTCCC		227
404	Qy	AACTGACCCGGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTGTCTCTCAAAACTAC	463	
228	Db	AACTGACCTGGGTACATAGCACCCCTGCGCCCTACAAAGGACTAGATCTCTCAAAACTAC		287
464	Qy	ATGAAACCCCTCGTACCCATACTCGCCTGGTGGAGCCTATTTAATACCAACCTTCATCTCGGC	523	
288	Db	ATGAAACCCCTCCATACCACCTACTGCGCCTGGTAAGCCTATTTAATACCAACCTTCATCTCGGC		347
524	Qy	TCCATGAGGTCTCAGGCCAAACCCCTACTAACTGTTGGATGTGCCTCCCCCTGCACCTTCA	583	
348	Db	TCCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATGTGCCTCCCCCTGCACCTTCA		407
584	Qy	GGCCATACATTTCAATCCCTGTTCTTGAAACAATGGAAACAACCTTCAGCA CAGAAATAACA	643	
408	Db	GGCCATACATTTCAATCCCTACTCTGAACCAATGGAAACAACCTTCAGCA CAGAAATAACA		467

PN WO9902696-A1.
 XX 21-JAN-1999.
 XX 06-JUL-1998; 98WO-FR001442.
 XX 07-JUL-1997; 97FR-00008815.
 XX (INMR) BIO MERIEUX.
 XX
 PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
 XX WPI; 1999-120897/10.
 XX
 PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
 PT disease, and abnormal or failed pregnancy.
 XX
 PS Claim 1; Page 56-58; 106pp; French.
 XX
 CC This sequence represents clone cl.24.4 of the human endogenous retrovirus
 CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
 CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
 CC dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility
 XX
 SQ Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 0 U; 3 Other;
 Query Match 75.4%; Score 1116.4; DB 2; Length 1948;
 Best Local Similarity 94.2%; Pred. No. 0;
 Matches 1159; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTATCATCTTTCTCTTACTGTTCTTCTTACCCCTTTTGGTCTCACT 60
 DB 719 ATGGCCCTCCCTATCATATTTTCTTACTGTTGTTTTCACCTCTTCACTCTCACT 778
 QY 61 GCACCCCTCAGTCTGTACACACAGTAGTCTCCCTTACCAGAGTTTCTATGAGA 120
 DB 779 GCACCCCTCAGTCTGTACACACAGTAGTCTCCCTTACCAGAGTTTCTATGAGA 838
 QY 121 ACGCGGCTCTCGAATATTGATGCCCCCATATATAGGAGTTTATCTAAGGAAACTCC 180
 DB 839 ATGCAGCGTCCCGAATATTGATGCCCCCATGTAAGGAGTCTTTGTAAGGAAACCCC 898
 QY 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 240
 DB 899, ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 958
 QY 241 CATGCAAACTACTATTATTGAGCAGGAGAAATGATTAATCTAGTGTCTGGAGGACTT 300
 DB 959 CATGCAAACTACTATTATTGAGCAGGAGAAATGATTAATCTAGTGTCTGGAGGACTT 1018
 QY 301 GGAGCCACTGTCTGTGACTTACTTCCACCATACAGTATGCTGATGGGGTGAATT 360
 DB 1019 GGAGTCACTGTCTGTGACTTACTTCCACCAAACTGATATGCTGATGGGGTGAATT 1078
 QY 361 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGGACAT 420
 DB 1079 CAAGATCAGGCAAGAGAAAAACAAGTAAAGAGTAAATCTCCCAACTGACCCCGGGTACAT 1138
 QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTGTCTCTCAAAACTATACATGAAACCCCTCCGTACC 480
 DB 1139 GGCACCTCTAGCCCTCAAAAGGACTAGTGTCTCTCAAAACTATACATGAAACCCCTCCGTACC 1198
 QY 481 CATACTGCGCTGGTGAAGCTATTTAATACCACTCTCGGCTCCATGAGGTCTCAGCC 540
 DB 1199 CATACTGCGCTGGTGAAGCTATTTAATACCACTCTCGGCTCCATGAGGTCTCAGCC 1258
 QY 541 CAAACCCCTACTAATGTTGGATGTGCTCCCTCCCTGCACTTCAGGCCATACATTTCAATC 600

DB 1259 CAAAACCCCTACTAATGTTGGATATGCTCCCTGAACTTCAGGCCATATGTTTCAATC 1318
 QY 601 CCTGTTCTGTAA CAATGGAACTTCAGCACAGAAATAAAACACACTTCGTTTGTAGTA 660
 DB 1319 CCTGTACCTGAAACAATGGAACTTCAGCACAGAAATAAAACACACTTCGTTTGTAGTA 1378
 QY 661 GGACCTCTGTTTCCAACTGGAATACCCATACCTCAAACTCAGCTGCTGCTAAATTT 720
 DB 1379 GGACCTCTGTTTCCAACTGGAATACCCATACCTCAAACTCAGCTGCTGCTAAATTT 1438
 QY 721 AGCAATACTATAGACACACACAGCTCCCAATCAGCTCAGGTGGGTAAACACCTCCCAACCA 780
 DB 1439 AGCAATACTATAGACACACACAGCTCCCAATCAGCTCAGGTGGGTAACTCCTCCCAACCA 1498
 QY 781 ATAGTCTGCTACCTCCCTCAGGAATATTTTTGTCTGTGTACCTCAGCTTATCATTTGTTG 840
 DB 1499 ATAGTCTGCTACCTCCCTCAGGAATATTTTTGTCTGTGTACCTCAGCTTATCATTTGTTG 1558
 QY 841 AATGGCTCTTCAGAACTATGCTTCTCTCTCATTTCTTAGTGGCCCTATGACCATCTAC 900
 DB 1559 AATGGCTCTTCAGAACTATGCTTCTCTCTCATTTCTTAGTGGCCCTATGACCATCTAC 1618
 QY 901 ACTGACAGAGTTTATACAACTATGCTGCTACCTTAAGCCCAACAAACAAAGAGTACCCATT 960
 DB 1619 ACTGACAGAGTTTATACAACTATGCTGCTATGCTATCTAAGCCCAACAAACAAAGAGTACCCATT 1678
 QY 961 CTTTCTTTTGTATCAGAGCAGGAGTGTAGGAGTCTAGGTAAGTCTGGCAATGGCAGTATC 1020
 DB 1679 CTTTCTTTTGTATAGGAGCAGGAGTGTAGGTAAGTCTAGGTAAGTCTGGCAATGGCAGTATC 1738
 QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGCATGGAACAG 1080
 DB 1739 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGCATGGAACAG 1798
 QY 1081 GTCACTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 DB 1799 GTGCGCGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1858
 QY 1141 CAAAATCGAAGAGCTTTAGACTTTGCTAACCGCAAAAGAGGGGAAACCTGTTTATTTTAA 1200
 DB 1859 CAAAATCGAAGAGCTTTAGACTTTGCTAACCGCTGAAAGAGGGGAAACCTGTTTATTTTAA 1918
 QY 1201 GGAGAAAGACGCTGTTTATTTATGTTAATCAA 1230
 DB 1919 GGGAAGAATGCTGTTTATTTATGTTAATCAA 1948

Search completed: February 21, 2005, 07:45:51
 Job time : 776.929 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:54:34 ; Search time 260.948 seconds
(without alignments)
9286.612 Million cell updates/sec

Title: US-09-319-156B-9
Perfect score: 1481
Sequence: 1 agggccctccctatcatcac.....gtctacaatggaaaccca 1481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	1481	4	US-08-979-847B-105
2	1335.4	90.2	2946	3	Sequence 105, App
3	1333.8	90.1	2763	4	Sequence 3, Appli
4	1333.8	90.1	8523	4	Sequence 5640, Ap
5	1333.8	90.1	13537	4	Sequence 21, Appl
6	1277.4	86.3	145320	4	Sequence 17382, A
7	302	20.4	1329	4	Sequence 15858, A
8	173.8	11.7	542	1	Sequence 108, App
9	173.8	11.7	542	1	Sequence 48, Appl
10	161.6	10.9	601	4	Sequence 2, Appli
11	160.8	10.9	190078	4	Sequence 184496
12	160.8	10.9	190078	4	Sequence 12707, A
13	150.8	10.2	601	4	Sequence 17026, A
14	150.8	10.2	601	4	Sequence 202250
15	150.8	10.2	77772	4	Sequence 202251
16	150.8	10.2	77997	4	Sequence 17417, A
17	145.6	9.8	1722	3	Sequence 12249, A
18	145.6	9.8	1722	3	Sequence 58, Appl
19	145.6	9.8	1722	4	Sequence 58, Appl
20	144.6	9.8	2908	4	Sequence 54, Appl
21	142.4	9.6	8399	4	Sequence 249, App
22	131.2	8.9	1859	3	Sequence 26, Appl
23	131.2	8.9	1859	4	Sequence 46, Appl
24	131.2	8.9	1859	4	Sequence 46, Appl
25	130.2	8.8	1001	4	Sequence 42, Appl
26	116.8	7.9	1704	1	Sequence 406, App
27	116.8	7.9	21526	4	Sequence 1, Appli
					Sequence 14685, A

c	28	116.6	7.9	601	4	US-09-949-016-184497	Sequence 184497,
	29	115.2	7.8	7713	4	US-09-573-080A-30	Sequence 30, Appl
	30	115	7.8	7392	4	US-09-573-080A-44	Sequence 44, Appl
c	31	114.8	7.8	601	4	US-09-949-016-184495	Sequence 184495,
	32	114.8	7.8	80246	3	US-09-078-294-4	Sequence 4, Appli
	33	114.2	7.7	2518	3	US-09-011-745-1	Sequence 1, Appli
	34	114.2	7.7	5865	3	US-09-011-745-8	Sequence 8, Appli
	35	112.6	7.6	6529	4	US-09-573-080A-28	Sequence 28, Appl
	36	112.2	7.6	4776	4	US-09-555-352-1	Sequence 1, Appli
	37	106.4	7.2	2809	4	US-09-620-312D-171	Sequence 171, App
	38	104.8	7.1	2342	4	US-09-904-615-12	Sequence 12, Appl
c	39	103.4	7.0	80355	4	US-09-949-016-12735	Sequence 12735, A
c	40	103.4	7.0	80357	4	US-09-949-016-13572	Sequence 13572, A
c	41	96.6	6.5	46698	4	US-09-949-016-17323	Sequence 17323, A
c	42	96.6	6.5	54382	4	US-09-949-016-12139	Sequence 12139, A
	43	90.8	6.1	80595	3	US-09-078-294-3	Sequence 3, Appli
	44	89	6.0	464	3	US-09-120-653D-6	Sequence 6, Appli
	45	89	6.0	3910	3	US-09-120-653D-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-979-847B-105
Sequence 105, Application US/08979847B
Patent No. 6582703
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
BESEME, FREDERIC
BEDIN, FREDERIC
PARAMHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUXE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-No. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-08-979-847B-105

Query Match 100.0%; Score 1481; DB 4; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACCTTTCTCTTACTGTTCTTCTACCCCTTTTCGCTCTCACT 60
DB 1 ATGGCCCTCCCTTATCATACCTTTCTCTTACTGTTCTTCTACCCCTTTTCGCTCTCACT 60

QY 61 GCACCCCTCCCTGCTGTGTACACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 61 GCACCCCTCCCTGCTGTGTACACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120

QY 121 AGCGGCTTCTGGAAATATTGATGCCCATATAGAGTTTATCTAAGGGAACCTCC 180
DB 121 AGCGGCTTCTGGAAATATTGATGCCCATATAGAGTTTATCTAAGGGAACCTCC 180

QY 181 ACCTTCACGCCCCACACCCATATGCCGCACTGCTATACTCTGCCACTCTTTGCAATG 240
DB 181 ACCTTCACGCCCCACACCCATATGCCGCACTGCTATACTCTGCCACTCTTTGCAATG 240

QY 241 CATGCAAACTACTTATTTGGACAGGGAATGATTAATCTCTAGTTGTCTGGAGACTT 300
DB 241 CATGCAAACTACTTATTTGGACAGGGAATGATTAATCTCTAGTTGTCTGGAGACTT 300

QY 301 GGAGCCACTGTCTGTTGGACTTACTTCCACCATACAGTATGCTGTGAGTGGGGTGAAT 360
DB 301 GGAGCCACTGTCTGTTGGACTTACTTCCACCATACAGTATGCTGTGAGTGGGGTGAAT 360

QY 361 CRAAGTTCAGGCAAGGAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
DB 361 CRAAGTTCAGGCAAGGAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420

QY 421 AGCACCCCTAGCCCTTACAAAGSACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
DB 421 AGCACCCCTAGCCCTTACAAAGSACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480

QY 481 CATACTCGCTGTGAGCCTATTATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB 481 CATACTCGCTGTGAGCCTATTATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540

QY 541 CAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCCTTCCAGGCCATACATTTCAATC 600
DB 541 CAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCCTTCCAGGCCATACATTTCAATC 600

QY 601 CTTGTTCTTGAAACAATGGAAACAATTCAGCACAGAAATAAACACCACTTCGGTTTAGTA 660
DB 601 CTTGTTCTTGAAACAATGGAAACAATTCAGCACAGAAATAAACACCACTTCGGTTTAGTA 660

QY 661 GGACCTCTGTTTCCAACTCGGAATAACCCATACCTCAAACTCACCTGTTGAAATTT 720
DB 661 GGACCTCTGTTTCCAACTCGGAATAACCCATACCTCAAACTCACCTGTTGAAATTT 720

QY 721 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGTGGGTAAACACCTCCACAGA 780
DB 721 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGTGGGTAAACACCTCCACAGA 780

QY 781 ATAGTCTGCTACCTCAGAAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 840
DB 781 ATAGTCTGCTACCTCAGAAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 840

QY 841 AATGGCTCTTCAGAACTATGCTTCTCTCAATCTTAGTGCCCTTATGACCACTCTAC 900
DB 841 AATGGCTCTTCAGAACTATGCTTCTCTCAATCTTAGTGCCCTTATGACCACTCTAC 900

QY 901 ACTGAACAAGATTTATACATCATGTCGTACCTAAGCCCCACAAAGAGTACCATT 960
DB 901 ACTGAACAAGATTTATACATCATGTCGTACCTAAGCCCCACAAAGAGTACCATT 960

QY 961 CTTCCCTTTGTTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGGCATGGCAGTATC 1020
DB 961 CTTCCCTTTGTTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGGCATGGCAGTATC 1020

RESULT 2

US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R. A.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B-AJ172A
; CURRENT APPLICATION NUMBER: US/09/175.928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3

Query Match 90.2%; Score 1335.4; DB 3; Length 2946;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACCTTTCTCTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
DB 928 ATGGCCCTCCCTTATCATATTTTCTCTTACTGTTCTTACCTCTTCACTCTCACT 987

QY 61 GCACCCCTCCATGCTGTGTACAAACCAAGTACCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 988 GCACCCCTCCATGCTGTGTATGACCAAGTACCTCCCTTACCAAGAGTTTCTATGGAGA 1047

QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC 180

Db	1123	CAAGATCAGGCAAGAGAAAAACAATGTATAAAGAGTAATCTCCCACTCACCCGGGTACAT	1187
Qy	421	AGCAACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAAACTACATGAACCCCTCGGTACC	480
Db	1183	GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC	1242
Qy	481	CATACTCGCCTGGTGAGCCTATTTTAATACACCCCTCACTCGGTCCATGAGGTCTCAGCC	540
Db	1243	CATACTCGCCTGGTAAGCCTATTTAATACACCCCTCACTGGGTCCATGAGGTCTCGGCC	1302
Qy	541	CAAAACCCCTACTAATCTTGATGTGCTCCCTCCCTGCACCTCAGGCGCATACATTTCAATC	600
Db	1303	CAAAACCCCTACTAATCTTGATGTGCTCCCTCCCTGCACCTCAGGCGCATATGTTTCAATC	1362
Qy	601	CCTGTTCTTGAACAAATGGAAACAATCTCAGCACAGAAATAAACACCACTTCGGTTTTAGTA	660
Db	1363	CCTGTACTGAACAATGGAAACAATCTCAGCACAGAAATAAACACCACTTCGGTTTTAGTA	1422
Qy	661	GGACCTCTGTTTCCCAATCTGGAAATAAACCACTACCTCAAACTCTGCTGTAAATTT	720
Db	1423	GGACCTCTGTTTCCCAATCTGGAAATAAACCACTACCTCAAACTCTGCTGTAAATTT	1482
Qy	721	AGCAATACTATPAGACAAACAGCCTCCCAATGCATCAGTGGGTAAACACCTCCACACGA	780
Db	1483	AGCAATACTATACAAACCACTCCCAATGCATCAGTGGGTAACTCCTCCACACAA	1542
Qy	781	ATAGTCTGCCTACCCCTCAGGAATATTTTTGTCGTGTACCTCAGCCTATCATTTGTTTG	840
Db	1543	ATAGTCTGCCTACCCCTCAGGAATATTTTTGTCGTGTACCTCAGCCTATCGTTGTTTG	1602
Qy	841	AATGGCTCTCAGAAATCTATGTGCTCTCTCTCATTTCTAGTGGCCCCCTATGACCATCTAC	900
Db	1603	AATGGCTCTCAGAAATCTATGTGCTCTCTCTCATTTCTAGTGGCCCCCTATGACCATCTAC	1662
Qy	901	ACTGAACAAGATTTATACAATCTGTGTGTAACCTTAAGGCCCAACAACCAAGTACCAATT	960
Db	1663	ACTGAACAAGATTTATACAGTTATGTGCATATCTAAGCCCGCGCAACAACCAAGTACCCATT	1722
Qy	961	CTTCTCTTTGTTATCAGACGAGGAGTGTCTAGGCAGACTAGGTACTGGCATTTGGCGGTATC	1020
Db	1723	CTTCTCTTTGTTATAGGACGAGGAGTGTCTAGGTGCACTAGGTACTGGCATTTGGCGGTATC	1782
Qy	1021	ACAACTCTACTCAGTTCTCTACAAACTATCTCAAGAAATAATGGTGATGCAATGGAACAG	1080
Db	1783	ACAACTCTACTCAGTTCTCTACAAACTATCTCAAGAAATAATGGTGATGCAATGGAACAG	1842
Qy	1081	GTCACGTACTCCCTGGTCACCTTCAAGATCAACTTAACCTCCCTAGCAGCAGTACTGCTT	1140
Db	1843	GTGCGCGACTCCCTGGTCACCTTGCAGATCAACTTAACCTCCCTAGCAGCAGTACTGCTT	1902
Qy	1141	CAAAATCGAAGAGCTTTAGACTTGCTATACCGCCAAAGAGGGGAACTGTTTATTTTTTA	1200
Db	1903	CAAAATCGAAGAGCTTTAGACTTGCTATACCGCTGAAGAGGGGAACTGTTTATTTTTTA	1962
Qy	1201	GGAGAAGAACGCTGTTATATGTGTTAATCAATCCGAATTTGTCACCTGAGAAAGTTTAAAGAA	1260
Db	1963	GGGGAAGAATGCTGTTATTTATGTTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA	2022
Qy	1261	ATTCGAGATCGAATAACAATGTAGACAGAGGAGCTTTCAAAACACCGCAACGCTGGGGCTC	1320
Db	2023	ATTCGAGATCGAATAACACGATAGACAGAGGAGCTTTCAAAACACCTGGGGCTC	2082
Qy	1321	CTCAGCCAATGATGCGCTGGGTTCTCCCTCTTTAGCACTCTAGCAGCTCTAATATTTG	1380
Db	2083	CTCAGCCAATGATGCGCTGGGTTCTCCCTCTTTAGCACTCTAGCAGCTCTAATATTTG	2142
Qy	1381	TTACTCTCTTTTGGACCCCTGTATCTTAACTCCCTTGTTTAGTTTGTCTCTTCCAGAAAT	1440
Db	2143	CTACTCTCTTTTGGACCCCTGTATCTTAACTCCCTTGTTTAGTTTGTCTCTTCCAGAAAT	2202
Qy	1441	GAAGCTGTAAGCTACAGATGTCCTTACAAATGGAAACCCCA	1491
Db	2203	GAAGCTGTAAGCTACAAATGGAGCCCAAGATGCGAGTCCAA	2243

RESULT 4

```

US-09-573-080A-21
; Sequence 21, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GEN
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/57
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: herV17
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichewicz, J;
; TITLE: Prototypic sequences for hum
; JOURNAL: Journal of Molecular Evolu
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10
; DATABASE ACCESSION NUMBER: Database
; DATABASE ENTRY DATE:
; DATABASE ENTRY DATE: 1996-01-26
; US-09-573-080A-21

```

Query Match	90.1%;	Score 1333.8;	DB 4;	Length 8523;	
Best Local Similarity	93.8%;	Pred. No. 0;			
Matches 1389;	Conservative	0;	Mismatches 92;	Indels 0;	Gaps 0;
QY	1	ATGGCCCTCCCTTATCATACATCTTCTCTTACTGTCTCTTACCCCTTTTCGTCTCACT	60		
Db	6849	ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTTACCTCTTTCACTCTCACT	6908		
QY	61	GCACCCCTCCAGTCTGTGACCAACAGTAGCTCCCTTACCAAGATTTCTTATGAAGA	120		
Db	6909	GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTTACCAAGATTTCTTATGAGA	6968		
QY	121	ACGCGGCTTCTCGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC	180		
Db	6969	ATGCAAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTCTTAAAGGAACCCCC	7028		
QY	181	ACCTTCACTGCCACACCCATATGCCCCCACTGCTATAACTCTGCCACTCTTTGGCATG	240		
Db	7029	ACCTTCACTGCCACACCCATATGCCCCCACTGCTATCACTCTGCCACTCTTTGGCATG	7088		
QY	241	CATGCAAACTACTCATTTATTGGACAGGGAATAATGATTAATCCTAGTTGTCTCTGGAGGACTT	300		
Db	7089	CATGCAAACTACTCATTTATTGGACAGGGAATAATGATTAATCCTAGTTGTCTCTGGAGGACTT	7148		
QY	301	GGAGCCACTGTCTGTTGGACTTATTCACCCCATACCAAGTATGCTGATGGGGTGGGAATT	360		
Db	7149	GGAGTCACTGTCTGTTGGACTTATTCACCCCAACTGGTATGCTGATGGGGTGGAGTT	7208		
QY	361	CAAGGTCAAGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT	420		
Db	7209	CAAGATCAAGCAAGAGAAAAACAATGTAAGGAAGTAACTCTCCAACTCAGCCGGGTACAT	7268		
QY	421	AGCACCCCTAGCCCTTACAAAGACTAGTCTCTCAAACTACATGAACCCCTCCGTACC	480		
Db	7269	GGCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC	7328		

Qy	721	AGCAATCTATAGACACAACACGAGCTCCCAATGCGATCAGGTGGGTAAACCTCCCAACGGA	780
Db	10256	AGCAATACTACATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCTCTCCCAACAA	10315
Qy	781	ATAGTCTGCCCTACCCCTCAGGNATATTTTTGTCGTGGTACCTCAGCCCTATCATTTGTTG	840
Db	10316	ATAGTCTGCCCTACCCCTCAGGAATATTTTTGTCGTGGTACCTCAGCCCTATCGTTGTTG	10375
Qy	841	AATGGCTCTTCCGAATCTATGTGCTTCCTCTCATTTAGTGGCCCCCTATGACCATCTAC	900
Db	10376	AATGGCTCTTCAGAACTATGTGCTTCCTCTCATTTAGTGGCCCCCTATGACCATCTAC	10435
Qy	901	ACTGAAACAAGATTTATACAAATCATGTCTACCTTAAGCCCCCAACAACAAGAGTACCCATT	960
Db	10436	ACTGAAACAAGATTTATACAGTTATGTGCATATCTAAGCCCCGCAACAACAAGAGTACCCATT	10495
Qy	961	CTTTCCTTTTGTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACTGGCATTTGGCAGATAC	1020
Db	10496	CTTTCCTTTTGTATAGGAGCAGGAGTGCTAGGTGCATCTAGGTACTGGCATTTGGCGGTATC	10555
Qy	1021	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGCAATGGAAACAG	1080
Db	10556	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGGGACATGGAAACGG	10615
Qy	1081	GTCACTGACTCCCTGTGTACCTTTGCAAGATCAACTTAATCCTCTAGCAGCAGTAGTCCTT	1140
Db	10616	GTCCGCACTCCCTGTGTACCTTTGCAAGATCAACTTAATCCTCTAGCAGCAGTAGTCCTT	10675
Qy	1141	CAAAATCGAAGAGCTTTAGACTTTGCTTAACCGCCAAAGAGGGGGAACTGTTTATTTTTA	1200
Db	10676	CAAAATCGAAGAGCTTTAGACTTTGCTTAACCGCTGAAAGAGGGGGAACTGTTTATTTTTA	10735
Qy	1201	GGAGAAACGCTGTTATTATGTTAATCAATCCAGAAATGTGCACATGAGAAAGTTAAAGAA	1260
Db	10736	GGGGAAGATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA	10795
Qy	1261	ATTCGAGATCGAATAACAATGTAGAGCAGAGGAGCTTCAAAAACACGAAACGCTGGGGCTC	1320
Db	10796	ATTCGAGATCGAATCAACGCTAGAGCAGAGGAGCTTCGAAAACACTGGACCCCTGGGGCTC	10855
Qy	1321	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
Db	10856	CTCAGCCAAATGGATGCCCTGGGATCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG	10915
Qy	1381	TTACTCTCTTTTGGACCCCTGATCTTTTAACCTCCTTGTAAAGTTTGTCTCTTCCAGAAAT	1440
Db	10916	CTACTCTCTTTTGGACCCCTGATCTTTTAACCTCCTTGTAACTTTGTCTCTTCCAGAAATC	10975
Qy	1441	GAAGCTCTAAGCTTACAGATGGTCTTCAAAATGGAAACCCCA	1481
Db	10976	GAAGCTCTAATAAATCAAAATGGAGCCCAAGATGCAGTCCAA	11016

```

RESULT 6
US-09-949-016-15858/c
; Sequence 15858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15858

```

; LENGTH: 145320									
; TYPE: DNA									
; ORGANISM: Human									
US-09-949-016-15858									
Query Match 86.3%; Score 1277.4; DB 4; Length 145320;									
Best Local Similarity 92.5%; Pred. No. 0;									
Matches 1399; Conservative 0; Mismatches 81; Indels 33; Gaps 4;									
Qy	1	ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTCTCTTAACCCCTTCGCTCTCACT	60						
Db	94027	ATGGCCCTCCCTTATCATATTTTCTCTTTACTGTCTCTTAACCCCTTCACCTCTCACT	93968						
Qy	61	GCACCCCTCCATGCTGCTGTACACAGTAGTCCCTTACCAAGAGTTCCTATGAGA	120						
Db	93967	GCACCCCTCCATGCTGCTGTACACAGTAGTCCCTTACCAAGAGTTCCTATGAGA	93908						
Qy	121	ACCGCGTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC	180						
Db	93907	ATCGGGCTTCCCGAANAATTTGATGCCCATCAATAGAGTTCCTTAAGGGAACCTCC	93848						
Qy	181	ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATAACTCTGCGCACTCTTTGCAATG	240						
Db	93847	ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATAACTCTGCGCACTCTTTGCAATG	93788						
Qy	241	CATGCAAACTACTATTATTGGACAGGGAATGATTATCTAGTTGTCTCTGGAGGACTT	300						
Db	93787	CATGCAAACTACTATTATTGGACAGGGAATGATTATCTAGTTGTCTCTGGAGGACTT	93728						
Qy	301	GGAGCCACTGCTGTTTGGACTTACTTTACCCATACAGTAGTCTCTGATGGGGTGGAAAT	360						
Db	93727	GGAGCCACTGCTGTTTGGACTTACTTTACCCATATCGTATGTCGAGGGGGTGGAGTT	93668						
Qy	361	CAAGGTGAGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCCAACTGACCCGGGACAT	420						
Db	93667	CAAGATCAGCAAGAGAAAAACATGTAAAGGAAGTAATCTCCCAACTGACCCGGGTACAT	93608						
Qy	421	AGCACCTTAGCCCTACAAAGGACTAGTTCTCTC-AAAACTACATGAAACCTCCGTTAC	479						
Db	93607	AGCACCTTAGCCCTACAAAGGACTAGTACTCTCTAATAACCTTACATGAACCCCTCCATAC	93548						
Qy	480	CCATCTCGCTGGTGAGCCTATTAAATACACCTCTACTCGGCTCCATGAGGTCTCAGC	539						
Db	93547	CCATCTCGCTGGTGAAGCCTATTAAATACACCTCTACTCGGCTCCATGAGGTCTCAGC	93488						
Qy	540	CCAAAACCTTACTAACTGTTGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAAT	599						
Db	93487	CCAAAACCTTACTAACTGTTGATGTGCTCCCTCGCTACTTCAGGCCATGCAATTTCAAT	93428						
Qy	600	CCCTGTTCTGAACATGGAACAACTTCAGCAGAGAAATAAACACCTTCGTTTTAGT	659						
Db	93427	CCCTGTTACCTGGAACAACTTCAGCAGCAGAAATAAACACCTTCGTTTTAGT	93368						
Qy	660	AGGACCTCTGTTTCCAAATCTCGAATAACCCATACCTCAAACTTCAGCTGTGTAATAAT	719						
Db	93367	AGGACCTCTGTTTCCAAATCTCGAATAACCCATACCTCAAACTTCAGCTGTGTAATAAT	93308						
Qy	720	TAGCAATACTATAGACACACAGCTCCCAATGCATCAGGTGGGTAAACACTCCACACG	779						
Db	93307	TAGCAATACTATAGACACACCAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACG	93248						
Qy	780	AATAGTCTGCCTACCTCAGGAATATTTTTTGTCTGTGTAGTACCTCAGCCTATCATGTTT	839						
Db	93247	AATAGTCTGCCTACCTCAGGAATATTTTTTGTCTGTGTAGTACCTCAGCCTATCATGTTT	93188						
Qy	840	GAATGGCTCTTCAGAACTCATGTGCTTCCTCTCATTTTAGTGCCCCCTATGACCATCTA	899						
Db	93187	GAATGGCTCTTCAGAACTCATGTGCTTCCTCTCATTTTAGTGCCCCCT-ATGACCATTTA	93129						
Qy	900	CACGAAACAGATTTTATACAAATCATGTCTGTAACCTAAGCCCCACACAAAGAGTACCCAT	959						
Db	93128	CACGAAACAGATTTTATACAAATATGTGTGTAACCTAAGCCCCACACAAAGAGTACTCAT	93069						

QY 960 TCTTCCTTTTGTATCAGACGAGGAGTCTAGCAGACTAGTAGTCTGCGATTGGCAGTAT 1019
Db 93068 TCTTCCTTTTGTATCAGACGAGGAGTCTAGTGAAGTCTGCGATTGGCAGTAT 93009
QY 1020 CACAACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATAATGGTGACATGAACA 1079
Db 93008 CACAACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATAATGGTGACATGAATG 92949
QY 1080 GGTCACTGACTCCTCGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTACTCT 1139
Db 92948 GGTGCGGACTCCTCGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTACTCT 92889
QY 1140 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTATTTT 1199
Db 92888 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTATTTT 92829
QY 1200 AGGAGAAGAACGCTG-----TTATATGTTTAATCA 1229
Db 92828 AGGGAAGAAATGTTTATTTAGCGGAAGAAATGTTTATTTATGTTAATCA 92769
QY 1230 ATCCAGAAATGTCACGTAGAGAAATTAAGAAATTCGAGATCGAATCAATGTAGAGCAGA 1289
Db 92768 ATCCAGAAATGTCACGTAGAGAAATTAAGAAATTCGAGATCGAATCAATGTAGAGCAGA 92709
QY 1290 GGAGCTTC-AAAACACCGAAGCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGGTTCTCC 1348
Db 92708 GGAGCTTC-AAAACACCGAAGCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGGTTCTCC 92649
QY 1349 CTTCTTTAGGACCTCTAGCAGCTCTAATATTTGTTTACTTCTCTTTGGACCTGTATCTTTA 1408
Db 92648 CTTCTTTAGGACCTCTAGCAGCTCTAATATTTGTTTACTTCTCTTTGGACCTGTATCTTTA 92589
QY 1409 ACTCCTTTGATGTTGTTCTCTCCAGAAATGAGCTGTAAGCTACAGATGGTCTTAC 1468
Db 92588 ACTCCTTTGATGTTGTTCTCTCCAGAAATGAGCTGTAAGCTACAGATGGTCTTAC 92529
QY 1469 AAATGGAAACCCA 1481
Db 92528 AAATGGAAACCCA 92516

RESULT 7
US-08-979-847B-108
; Sequence 108, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-No. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-08-979-847B-108
Query Match 20.4%; Score 302; DB 4; Length 1329;
Best Local Similarity 92.7%; Pred. No. 2.4e-87;
Matches 317; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1140 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTATTTT 1199
Db 1 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTATTTT 60
QY 1200 AGGAGAAGAACGCTGTTATTTATGTTAATCAATCCAGAAATTTGTCAGTGAAGATTAAGA 1259
Db 61 AGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
QY 1260 AATTCAGATCGAATCAATGTAGACAGAGGAGCTTCAAAACACCGAAGCTGGGGCCT 1319
Db 121 AATTCAGATCGAATCAATGTAGACAGAGGAGCTTCAAAACACTGCACCTGGGGCCT 180
QY 1320 CCTCAGCAATGATGCCCTGGGTTCTCCCTTTCTAGGACCTCTAGCAGCTCTAATATT 1379
Db 181 CCTCAGCAATGATGCCCTGGGTTCTCCCTTTCTAGGACCTCTAGCAGCTCTAATATT 240
QY 1380 GTTACTCTCTTTGGACCTGTTATCTTTAACTCTCTTTGTTAGTTGTTCTCTCCAGAA 1439
Db 241 TTTACTCTCTTTGGACCTGTTATCTTTCACTCTCTTTGTTAGTTGTTCTCTCCAGAA 300
QY 1440 TGAAGCTGTAAGCTACAGATGCTTTACAAATGGAACCCCA 1481
Db 301 TGAAGCTGTAAGCTACAGATGCTTTCTTCAATGGAACCCCA 342

RESULT 8
US-08-686-878A-48
; Sequence 48, Application US/08686878A
; Patent No. 5708157
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Bridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 541 CAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGCCTTACAGGCATATCTTCAATC 600
 DB 412 GGTAAATCCAACTGCTGGATGCTCCCTTGGCA-TTTTCAATGATGTCCTCCAGTC 354
 QY 601 CTGTGCTTCTGAACTGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTA 660
 DB 353 CTGTGCTTCTGAACTGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTA 660
 QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCTTCAACCTTCAACCTTCAACCTTCAACCTT 720
 DB 296 AGTCCCATAGTACCAATTTTACCAGCCACACAGGCTTCAATCTCAGCTGATCAAACTTC 237
 QY 721 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
 DB 236 AGCATGCTGTCAATAGAACACCTTCTGATGATGATGATGATGATGATGATGATGATG 177
 QY 781 ATAGTCTGCTTACCTTCCAGGAATATTTTGTCTGTGTGATGATGATGATGATGATGATG 840
 DB 176 TTCACTTAACTTTCAGGCA--TTTGTCTGTGTGATGATGATGATGATGATGATGATG 119
 QY 841 AATGGCTCTTCAAGATCTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
 DB 118 AACAGCACTCTGAAAGAACTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 59
 QY 901 ACTGAACAGATTTATACATCATGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 952
 DB 58 ACTGAACAGATTTATACATCATGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7

RESULT 11

US-09-949-016-12707
 ; Sequence 12707, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12707
 ; LENGTH: 190078
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(190078)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-12707

Query Match 10.9%; Score 160.8; DB 4; Length 190078;
 Best Local Similarity 59.8%; Pred. No. 5.2e-40;
 Matches 324; Conservative 0; Mismatches 212; Indels 6; Gaps 3;
 QY 421 AGCACCCCTAGCCCTCAAGAGACTAGTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 DB 95515 AGTACTCCAGTCCATCAAGAAATAGACCTTTCCAGGCTACAGGAAGCCCTTAACTCT 95574
 QY 481 CATACTCGCTGGTGGCTATTTATACCACTCTACTCGGCTCCATGAGTCTCAGCC 540
 DB 95575 CATTTCTATCTCGAGGCTCTTTAACACCACTTACAGGAATACAGGAAGCCCTCTCCT 95634
 QY 541 CAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGCCTTACAGGCATATCTTCAATC 600

DB 95635 GGTAAATCCAACTGCTGGATGCTCCCTTGGCA-TTTTCAATGATGATGATGATGATGATG 95693
 QY 601 CTGTGCTTCTGAACTGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTA 660
 DB 95694 CTGTGCTTCTGAACTGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTA 660
 QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCTTCAACCTTCAACCTTCAACCTTCAACCTT 720
 DB 95751 AGTCCCATAGTACCAATTTTACCAGCCACACAGGCTTCAATCTCAGCTGATCAAACTTC 95810
 QY 721 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
 DB 95811 AGCATGCTGTCAATAGAACACCTTCTGATGATGATGATGATGATGATGATGATGATG 95870
 QY 781 ATAGTCTGCTTACCTTCCAGGAATATTTTGTCTGTGTGATGATGATGATGATGATGATG 840
 DB 95871 TTCACTTAACTTTCAGGCA--TTTGTCTGTGTGATGATGATGATGATGATGATGATG 95928
 QY 841 AATGGCTCTTCAAGATCTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
 DB 95929 AACAGCACTCTGAAAGAACTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 95988
 QY 901 ACTGAACAGATTTATACATCATGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
 DB 95989 ACTGAACAGATTTATACATCATGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 96048
 QY 961 CT 962
 DB 96049 GT 96050

RESULT 12

US-09-949-016-17026
 ; Sequence 17026, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17026
 ; LENGTH: 190078
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(190078)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17026

Query Match 10.9%; Score 160.8; DB 4; Length 190078;
 Best Local Similarity 59.8%; Pred. No. 5.2e-40;
 Matches 324; Conservative 0; Mismatches 212; Indels 6; Gaps 3;
 QY 421 AGCACCCCTAGCCCTCAAGAGACTAGTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 DB 95515 AGTACTCCAGTCCATCAAGAAATAGACCTTTCCAGGCTACAGGAAGCCCTTAACTCT 95574
 QY 481 CATACTCGCTGGTGGCTATTTATACCACTCTACTCGGCTCCATGAGTCTCAGCC 540
 DB 95575 CATTTCTATCTCGAGGCTCTTTAACACCACTTACAGGAATACAGGAAGCCCTCTCCT 95634
 QY 541 CAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGCCTTACAGGCATATCTTCAATC 600

Db 95635 GGTAAATCAACCAACTGCTGGATGTGTCTCCCTTTGCA-TTTTCAATGATATGTCCCACTC 95693
Qy 601 CCGTGTTCCTGAACAAATGGAACAATTCAGCAGAGAAATAAACACCACTTCGGTTTATGTA 660
Db 95694 CCGTGTTCCTGGACAGTGGAACTTATCCACCAGAGTACTAAACATCA---CCAAATCAATC 95750
Qy 661 GGACCTCTTGTGTTCCTGGAATAACCCATACCTCAAACTCAGCTGTGTGTAATAATTT 720
Db 95751 AGTCCCATAGTCACCAATTTTACCAGCCACACAGGCTCAAACTCTCACGTGCAATAACTTC 95810
Qy 721 AGCAATACATAGACACACACAGCTCCCAATCATCAGGTGGTAACACCTCCCAACGCA 780
Db 95811 AGCATGGCTGTCAATAAGAACACCTTCTGATCTAGCTGTGATATCATAGTAACCCAGGT 95870
Qy 781 ATAGTCTGCTACCTCCAGCAATATTTTGTCTGTGGTACCTCAGCCCTATCATTTGTTG 840
Db 95871 TTCACCTAATCTCAGCA--TTTTTTTTTTCATCTGTGATTACACAGGTTATGATGCCTA 95928
Qy 841 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCCGCCCTATGACCATCTAC 900
Db 95929 AACAGCACTCTGAAGAACTATGCTTCTCACTCCTTTCTAGCACCTCCCATGTCCATATAT 95988
Qy 901 ACTGAACAAGTTTATACATCATGTCGTACCTTAAGCCCAACACAAAGAGGTACCCATT 960
Db 95989 ACTGAACGAGGTTACAAAGTCTCCTTATACCCCAATCTCGCCACACATGAGCCTTTATT 96048
Qy 961 CT 962
Db 96049 GT 96050

RESULT 13
US-09-949-016-202250
; Sequence 202250, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 202250
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202250

Query Match 10.2%; Score 150.8; DB 4; Length 601;
Best Local Similarity 84.6%; Pred. No. 2.5e-38;
Matches 181; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
Qy 4 GCCCTCCCTTATCATCTTTCTTTACTGTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 62
Db 297 GGCWCCCCCGCTATATTTTCTTTACTGTGTCTTACCCCTTTCATATCACTACCTC 356
Qy 63 ACCCCCTCCATGCTGTGTAACAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGAAC 122
Db 357 ACCCCCTCCATGCGCTATACTACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGGAGAAT 416
Qy 123 GCGGCTTCCGGAATATTGATGCCCCCATATAGGAGTTTATCTAAGGGAACCTCCAC 182
Db 417 GCGGCTTCCGGAATATTGATGCCCCCATATGATAGGAGTTTCTTAAAGGAACCCCA 476

Qy 183 CTTCACTGCCACCAACCATATATGCCCGCAACTGC 216
Db 477 TTTCAACATCCACCAACCATATATGCCCTGCACTTC 510
RESULT 14
US-09-949-016-202251
; Sequence 202251, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 202251
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202251

Query Match 10.2%; Score 150.8; DB 4; Length 601;
Best Local Similarity 84.6%; Pred. No. 2.5e-38;
Matches 181; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
Qy 4 GCCCTCCCTTATCATCTTTCTTTACTGTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 62
Db 299 GGCWCCCCCGCTATATTTTCTTTACTGTGTCTTACCCCTTTCATATCACTACCTC 358
Qy 63 ACCCCCTCCATGCTGTGTAACAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGAAC 122
Db 359 ACCCCCTCCATGCGCTATACTACCAAGTAGTCCCTTACCAAGAGTTTCTATGGAGAAT 418
Qy 123 GCGGCTTCCGGAATATTGATGCCCCCATATAGGAGTTTATCTAAGGGAACCTCCAC 182
Db 419 GCGGCTTCCGGAATATTGATGCCCCCATATGATAGGAGTTTCTTAAAGGAACCCCA 478
Qy 183 CTTCACTGCCACCAACCATATATGCCCGCAACTGC 216
Db 479 TTTCAACATCCACCAACCATATATGCCCTGCACTTC 512

RESULT 15
US-09-949-016-17417/c
; Sequence 17417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17417
; LENGTH: 7772
; TYPE: DNA

```

RESULT 16
US-09-949-016-12249/c
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

```

```

183 CTTCACTGCCACACCCCATATGCCCGCACTGC 216
Db      16027 TTTCACCATCCACACCCATATGCCCTGCACCTTC 15994

RESULT 17
US-08-691-563C-58
; Sequence 58, Application US/08691563C
; Patent No. 6001987
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Frederic BESEME
; APPLICANT: Frederic BEDIN
; APPLICANT: Glaucia PARANHOS-BACCALA
; APPLICANT: Florence KOMURIAN-PRADEL
; APPLICANT: Collette JOLIVET
; APPLICANT: Bernard WANDRAND
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SC
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.300
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,563C
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1722 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-691-563C-58

```

Query Match	9.8%	Score 145.6;	DB 3;	Length 1722;
Best Local Similarity	55.7%;	Pred. No. 2.4e-36;		
Matches 302;	Conservative	0;	Mismatches 234;	Indels 6; Gaps 1;
Qy	940	CACAAACAAAAGAGTACCCATCTCTCTTTTGGTTTATCAGACGACGAGTGCTAGGCAGACTA	999	
Db	346	CAGTGCAGAGCCATACAACTAATATCCCTATTATTAGGTTAGGAATGGCTACTGTCTACA	405	
Qy	1000	GGTACTGGCAATTGGCAGTATTCACAACTCTACTCTAGTTCTACTACAAACTATCTCAAGAA	1059	
Db	406	GGAACTGGAAATAGCCGGTTTATCTACTTCAATATCCTACTACCATACACTCTCAAGAAT	465	
Qy	1060	ATAAATGGTGACATGGAACAGGTCACGTACTCCTCGTCTACCTTGAAGAATCAACTTTAAC	1119	
Db	466	TTCTCAGACAGTTTGCAGAAATTAATGAAATCTATTCTTACTTTACAAATCCCAATTAGAC	525	
Qy	1120	TCCTTAGCAGCAGTAGTCTCTTCAAATTCGAAGAGCTTTTAGACTTGTCACCGCCAAAGA	1179	
Db	526	TCTTTGGCAGCAATGACTCTCCAAAAACCGCGAGGCCACACCTCTCTACTGTGTGAGAA	585	
Qy	1180	GGGGCAACCTGGTTTTATTTTTTAGCAGAGAAGCGCTGTTATTATTGTGTTAATCAATCCAGAATT	1239	

Db 586 GGAGAGACTCGACCTCTTACGGGAAGAGTGTGTTTACCTAACCAAGTCAGGGATA 645
Qy 1240 GTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATAGAGAGAGAGGCTTCAA 1299
Db 646 GTACGAGATGCCACCTGGCATTTACAGGAAGGGCTTCTGTATATCAGACAATGCGCTTTCA 705
Qy 1300 AACACCGAAGCTGGGGCCCTCTCAGCCAAATCGATGCCCTGGTCTCCCTTCTTAGGA 1359
Db 706 AACTCTTATACCAA-----CCTCTGGAGTTGGCAACATGGCTTCTTCCATTTCTAGGT 759
Qy 1360 CCTCTAGCAGCTCTAATATGTTTACTCTCTTTGGACCCCTGTATCTTTAAACCTCCTTGT 1419
Db 760 CCATGGCAGCATCTTGCTGTACTACCTTTGGGCCCTGTATTTTAAAGCTTCTTGTC 819
Qy 1420 AAGTTGTCTCTTCAGAAATGAAGCTGTAAGAGCTACAGATGGCTTACAAATGGAACCC 1479
Db 820 AAATTTGTTCTCTAGGATCGAAGCCATCAAGCTACAGATGGTCTTACAAATGGAACCC 879
Qy 1480 CA 1481
Db 880 CA 881

RESULT 18
US-09-374-766-58
; Sequence 58, Application US/09374766
; Patent No. 6579526
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Frederic BESEME
; APPLICANT: Frederic BEDIN
; APPLICANT: Glauclia PARANHOS-BACCALA
; APPLICANT: Florence KOMURIAN-PRADEL
; APPLICANT: Collette JOLIVET
; APPLICANT: Bernard MANDRAND
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/374,766
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,563
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1722 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-374-766-58

Query Match 9.8%; Score 145.6; DB 4; Length 1722;
Best Local Similarity 55.7%; Pred. No. 2.4e-36;
Matches 302; Conservative 0; Mismatches 234; Indels 6; Gaps 1;
Qy 940 CACAACAAAAGAGTACCCATTCCTCTTTGTTATCAGAGCAGGAGTGCTAGGCGAGCTA 999
Db 346 CAGTGCAGAGCATACAACTAATATCCCTATTTATAGGGTTAGGAATGGCTACTGCTACA 405
Qy 1000 GGTACTGGCATTCGCGAGTATCACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAA 1059
Db 406 GGAATCTGGAATAGCGGTTTATCTACTTCTATATTCCTACTACCATACACTCTCAAGAAAT 465
Qy 1060 ATAAATGTGACATGGAACAGGTCACTGACTCCCTGTGTCACCTTTCGCAAGATCAACTTAAC 1119
Db 466 TTCTCAGACAGTTTGCAGAAATATGAAATCTATCTTACTTTACAATCCCAATTAGAC 525
Qy 1120 TCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGCTTAACCGCCAAAGA 1179
Db 526 TCTTTGGCAGCAATGACTCTCCAAAACCGCCGAGGCCACACCTCTCACTGCTGAGAAA 585
Qy 1180 GGGGGAACCTGTTTATTTTAGGAGAAAGACGCTGTTATTTATGTTTAAATCCAGAATT 1239
Db 586 GGAGGACTCTGCACCTTCTTAGGGGAAGAGTGTGTTTTTACCTAACCAAGTCAGGGATA 645
Qy 1240 GTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGGAGCTTCAA 1299
Db 646 GTACGAGATGCCACCTGCGATTTACAGAAAGGGCTTCTGATATCAGACAATGCGCTTTCA 705
Qy 1300 AACACCGAAGCTGGGGCCCTCTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGA 1359
Db 706 AACTCTTATACCAA-----CCTCTGGAGTTGGGCAACATGGCTTCTTCCATTTCTAGGT 759
Qy 1360 CCTCTAGCAGCTCTAATATGTTTACTCTCTTTGGACCCCTGTATCTTTAAACCTCCTTGT 1419
Db 760 CCCATGGCAGCATCTTGCTGTACTACCTTTGGGCCCTGTATTTTAAAGCTTCTTGTC 819
Qy 1420 AAGTTGTCTCTTCAGAAATGAAGCTGTAAGAGCTACAGATGGCTTACAAATGGAACCC 1479
Db 820 AAATTTGTTCTCTAGGATCGAAGCCATCAAGCTACAGATGGTCTTACAAATGGAACCC 879
Qy 1480 CA 1481
Db 880 CA 881

RESULT 19
US-08-979-847B-54
; Sequence 54, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKER, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

QY 1216 TATTATCTTAATCAATCAGAAATGTCTACGTGAGAAAGTTAAAGAAATTCGAGATCGAATA 1275
| | | | |
DB 1272 CTTTACACTAACAGTCAGGAGTAGATGATGCTGCCCGCATTTTACAGAAAAAGGCT 1331
| | | | |
QY 1276 CAATGTAGACGAGAGAGCTTAAACACCGAAGCTGGGGCTCTCTCAGCCCAATGGATG 1335
| | | | |
DB 1332 TCTGAATCAGACAAACGCTTTCAAAATTC-----CTATACCAACCTCTGGAGTTGGGCA 1385
| | | | |
QY 1336 CCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATTTTGTACTCTCTTTGGA 1395
| | | | |
DB 1386 ACATGGGTTCTTCCCTTTCTATGTCCTATGCTGCGCTGCTGCTTACTCTGCTTTGGG 1445
| | | | |
QY 1396 CCCTGTATCTTTAACTCTTTGTTAGTTGTTCTCTTCCAGAAATGCAAGCTGTAAAGCTA 1455
| | | | |
DB 1446 CCCTGTATTTTAACTCTCTTGTCATATTTGTTCTTCTAGGATCGAGGCCATCAAGCTA 1505
| | | | |
QY 1456 CAGATGCTTTACAAATGGAACCCCA 1481
| | | | |
DB 1506 CAGATGCTTTACAAATGGAACCCCA 1531
| | | | |

RESULT 23
US-09-374-766-46
; Sequence 46, Application US/09374766
; Patent No. 6579526
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Frederic BESEME
; APPLICANT: Frederic BEDIN
; APPLICANT: Glauicia PARANHOS-BACCALA
; APPLICANT: Florence KOMURIAN-PRADEL
; APPLICANT: Colette JOLIVET
; APPLICANT: Bernard MANDRAND
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/374,766
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,563
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1859 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-374-766-46
Query Match 8.9%; Score 131.2; DB 4; Length 1859;
Best Local Similarity 57.6%; Pred. No. 1.2e-31;

Matches 257; Conservative 0; Mismatches 183; Indels 6; Gaps 1;
QY 1036 TTCTACTACAAACTATCTCAAGAAATAAATGTGTGACATGGAACAGGTCTACTGCTCCCTG 1095
| | | | |
DB 1092 TACTACCACACTCTCAAAGGATTTCTCAGACAGTTTGCAGAAATAATGATATCTATC 1151
| | | | |
QY 1096 GTCACCTTCGAGATCACTTAATCCCTTAGCAGCAGTAGTCTTCAAAATCGAAGAGCT 1155
| | | | |
DB 1152 CTTACTCTACAATCCCAATAGACTCTTTGGCAGCAGTAGTCTTCCAAAACCGTCAAGGC 1211
| | | | |
QY 1156 TTAGACTTCTACCGCCCAAGAGAGGGGAACCTGTTTATTTTAGGAGAAGAACCTGT 1215
| | | | |
DB 1212 CTAGACCTCTCTCAGTCTGAGAAAGAGGAGCTCTGCACCTTCTTAAGGGAAGAGTGTGT 1271
| | | | |
QY 1216 TATTATGTAAATCAATCCAGAAATTTGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATA 1275
| | | | |
DB 1272 CTTTACACTAACAGTCAGGATAGTATGAGATGCTGCCCGCATTTACAGAAAAAGGCT 1331
| | | | |
QY 1276 CAATGTAGACGAGAGAGCTTCAAAACACCGAAGCTGGGGCTCTCTCAGCCCAATGGATG 1335
| | | | |
DB 1332 TCTGAAATCAGACACGCTTTCAAAATTC-----CTATACCAACCTCTGGAGTTGGGCA 1385
| | | | |
QY 1336 CCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATTTTGTACTCTCTTTGGA 1395
| | | | |
DB 1386 ACATGGGTTCTTCCCTTTCTATGTCCTATGCTGCGCTGCTGCTTACTCTGCTTTGGG 1445
| | | | |
QY 1396 CCCTGTATCTTTAACTCTTTGTTAGTTGTTCTCTTCCAGAAATGCAAGCTGTAAAGCTA 1455
| | | | |
DB 1446 CCCTGTATTTTAACTCTCTTGTCATATTTGTTCTTCTAGGATCGAGGCCATCAAGCTA 1505
| | | | |
QY 1456 CAGATGCTTTACAAATGGAACCCCA 1481
| | | | |
DB 1506 CAGATGCTTTACAAATGGAACCCCA 1531
| | | | |

RESULT 24
US-08-979-847B-42
; Sequence 42, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REVNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUBE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1859 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-08-979-847B-42

Query Match      8.9%; Score 131.2; DB 4; Length 1859;
Best Local Similarity 57.6%; Pred. No. 1.2e-31;
Matches 257; Conservative 0; Mismatches 183; Indels 6; Gaps 1;

QY 1036 TTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACCTCCCTG 1095
Db 1092 TACTACACACACTCTCAAGGATTTCTCAGACAGTTTGCAAGAAATATGATATCTATC 1151

QY 1096 GTCACTTCAGAGATCAACTTAATCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCT 1155
Db 1152 CTTACTCTCAATCCCAATAGACTCTTTGGCAGCAGTGACTCTCCAAAACCGTCAAGGC 1211

QY 1156 TTAGACTTGTAAACCCCAAGAGGGGAACCTGTTTATTATTTTAGGAGAAGAACGCTGT 1215
Db 1212 CTAGACCTCTCACTCTGAGAAAGGAGGACTCTGCACCTCTTAAAGGAAGAGTGTGT 1271

QY 1216 TATTATGTTAATCAATCCAGAAATTTGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATA 1275
Db 1272 CTTTACACTAACCGTCAGGATAGTATGAGATGCTGCCGGCATTTTACAGAAAAGGCT 1331

QY 1276 CAATGTAGCAGAGAGAGCTTCAAAACACCGAAGCGTGGGGCTCTCTAGCCAAATGGATG 1335
Db 1332 TCTGAAATCAGACAAGCGCTTTCAAAATC-----CTATACCAACCTCTGGAGTTGGGCA 1385

QY 1336 CCCTGGGTTCTCCCTCTTAGACCTCTAGACGCTCTAATATTGTTACTCTCTTTTGA 1395
Db 1386 ACATGGTTTCTCCCTTTCTATGTCCTCATGGCTGCGCATCTTGCTATTATCTCGCCTTTGG 1445

QY 1396 CCCTGTATCTTTAACTCTTTTAAAGTTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 1455
Db 1446 CCCTGTATTTTAACTCTCTGTCAATTTGTTCTTCTAGGATCGAGGCCATCAAGCTA 1505

QY 1456 CAGATGGCTTACAAATGGAACCCCA 1481
Db 1506 CAGATGGCTTACAAATGGAACCCCA 1531

RESULT 25
US-09-671-317-406
; Sequence 406, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
; FILE REFERENCE: 62 US3 CIP
; CURRENT APPLICATION NUMBER: US/09/671,317
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 09/536,178
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/IB00/00403
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 60/126,269
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/131,961
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 977
; SOFTWARE: Patent.pm

; SEQ ID NO 406
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 503
; OTHER INFORMATION: 12-5-158 : polymorphic base C or T
; NAME/KEY: misc binding
; LOCATION: 484..502
; OTHER INFORMATION: 12-5-158.mis1
; NAME/KEY: misc binding
; LOCATION: 504..523
; OTHER INFORMATION: 12-5-158.mis2, potential complement
; NAME/KEY: primer bind
; LOCATION: 346..366
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer bind
; LOCATION: 801..821
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 491..515
; OTHER INFORMATION: 12-5-158 potential probe
; NAME/KEY: misc feature
; LOCATION: 314,336,793
; OTHER INFORMATION: n=a, g, c or t
US-09-671-317-406

Query Match      8.8%; Score 130.2; DB 4; Length 1001;
Best Local Similarity 67.6%; Pred. No. 1.8e-31;
Matches 211; Conservative 0; Mismatches 99; Indels 2; Gaps 2;

QY 773 CCACGAGAAATAGTCTGCTACCTCAGG-AATATTTTGTGTGTGCTCAGCCTAT 831
Db 688 CCTCAGGTTTCACTATCTACCTCAGGAAATTTTTCATCTGTGATACACAGCCTAT 747

QY 832 CATGTTTGAATGGCTTTTCAGAAATCTATGTGCTTCCTCATCTT-TAGTGCCCCCTAT 890
Db 748 CGATGCCAAATGGCACTCCAAAGAACTATGTCTATGTCTCATTTCTAGCACCTCCCAT 807

QY 891 GACCATCTACACTGAAAGAATTTATCAATCATGTCTGCTAGTACGAGCCCCACAAAGAAG 950
Db 808 GTCCATATATATCTGAACAAGAGTTACAAAGTCTCCTTATACCCCAATCTCGCCACACCCG 867

QY 951 AGTACCCCATCTCTCTTTTGTATCAGAGCAGGAGTGTAGGACAGCTAGTACTTGGCAT 1010
Db 868 AGCCCTTATTTGCCCTTTTACAGTAGGAGCAGGATATCTAGGAGGCTTGGGACTGGAAT 927

QY 1011 TGGCAGTATCAAACTCTCTACTCTACTCTACTACAACTATCTACAACTATCTCAAGAAATAAATGGTGA 1070
Db 928 TGGAGGCATAACCTCCTCCACCAATTTCTATTATAAATTATCATGAGAATTAAATGATGA 987

QY 1071 CATGGAACAGGT 1082
Db 988 CATGGAATCAGT 999

RESULT 26
US-08-007-282B-1
; Sequence 1, Application US/08007282B
; Patent No. 5403582
; GENERAL INFORMATION:
; APPLICANT: NAZERIAN, KEYVAN
; APPLICANT: CALVERT, JAY G.
; APPLICANT: WITTER, RICHARD L.
; APPLICANT: YANAGIDA, NOBORU
; TITLE OF INVENTION: VACCINE COMPRISING FOWLPOX VIRUS
; TITLE OF INVENTION: RECOMBINANTS EXPRESSING THE ENVELOPE GLYCOPROTEIN OF AN
; TITLE OF INVENTION: AVIAN RETICULOENDOTHELIOSIS RETROVIRUS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road Suite 500 East
```

CITY: Falls Church
STATE: VA
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/007,282B
FILING DATE: 19930121
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR., GERALD M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1644-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-007-282B-1

Query Match	7.9%	Score 116.8;	DB 1;	Length 1704;
Best local Similarity	51.5%;	Pred. No. 5.8e-27;		
Matches 268;	Conservative 0;	Mismatches 252;	Indels 0;	Gaps 0;
Qy	932	CTAAGCCCCACAA	AAAAGAGTACCC	CAATCTTCTCTTTTGTATCAGACGAGGAGTGTAG 991
Db	1178	CTAGGGCTCAT	NAGAGGGCAGTCC	AGTTTATCCCTTCTGTTGGGTCTAGGATTTTCAG 1237
Qy	992	GCAGACTAGGT	ACTTGGCAGTATCA	CAACCTCTACTCTAGTTCTACTTACAAACTAT 1051
Db	1238	GGGCTACACT	TGCTGGTGGAA	CGGGCTCCGGTTCACACTTATCACAAGCTCT 1297
Qy	1052	CTCAGAAATAA	TGTTGACATGGA	ACAGGTCACTGACTCCCTGGTCACCTTGCAGATC 1111
Db	1298	CTAATCAATT	GAATTGAAGATGT	CCAGGCTCTTTCAGGGACCATCAATGACCTACAGACC 1357
Qy	1112	AACTTAACTCC	CTAGCAGCAGTAG	CTCTTCAAAATCGAAGAGCTTTAGACTTGTAAACCG 1171
Db	1358	AGATTGACTCC	TGCGCTGAGTTGT	CTTACAAATAGAGAGGCTTAGACCTATTGACTG 1417
Qy	1172	CCAAAAGAGGG	GAACCTGTTTAT	TTTTTAGAGAGAAGACGCTGTTATTATGTTAAATCAAT 1231
Db	1418	CCGAAACAAG	GAGGAATATGT	CTCGCACTCCAGGAGAAAGTGTGTTTTCGCTAAACAAGT 1477
Qy	1232	CCAGAAATTGT	CACTGAGAAAGT	TAAAGAAATTCAGATCGAATACAAATGTAGACAGAGG 1291
Db	1478	CGGGTATTCGT	ACGTGACAAGAT	CCGAAATCCGAAACCTCCAAGAGGACCTTTATCGAGAGAAAACGTG 1537
Qy	1292	AGCTTCAAAAC	CACCGAAGCGCT	GCGGCCTCTTCACGCAATGGATGCCCTGGGTCTCTCCCT 1351
Db	1538	CACGTGAGCA	CAACCCCTGTGG	AGCGGCTTGAACGGCTTCCTTCCATATTGTGTACCT 1597
Qy	1352	TCTTTAGGAC	CTCTAGCAGCTCA	ATAATTGTTATCTCTCTTTTGGACCTGTATCTTTAACC 1411
Db	1598	TGTTTAGCCCC	CTCTTTGGGCTCA	TATTGTTCTCTGACCTCGGCCCGTGCAATTATGAAGA 1657
Qy	1412	TCCTTGTTTA	GTTTGTCTCTTCC	AGAAATTTGAAGCTGTATAA 1451
Db	1658	CCCTGACTCG	ATATATACATGAC	AAAAATTCAGGCAGTAAA 1697

RESULT 27

US-09-949-016-14685/C

; Sequence 14685, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 14685

; LENGTH: 21526

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-14685

Query Match	7.9%	Score 116.8	DB 4	Length 21526	
Best Local Similarity	54.2%	Pred. No. 2.8e-26			
Matches 260	Conservative	0	Mismatches 217	Indels 3	Gaps 1
Qy	997	CTAGGTACTGGCAATGGCAGATACACAACCTCTACTCTCAGTTCAGTCTTCTACTACAACTATCTCAA	1056		
Db	9564	CTCGTACTGGAATAGGAGGCAITTCAAACCTGTGCAGACCTTCCATAGCCTCTCTAGT	9505		
Qy	1057	GAATAAATGGTGACATGGNAACAGGTCACTGACTCCTGGTCACTTGGCAATCAACTT	1116		
Db	9504	GTCTTCTCTCTAGCATACAGACATATCAAACTTTATCAGTCTCTCCAGGCCCAAGTT	9445		
Qy	1117	AACTCCCTACGACGATGTCCTTCAAATTCGAAGAGCTTTTAGACTTGTCTAAACCGCCAA	1176		
Db	9444	GACTCTTTAGTCGAGTTATCTCCAAACCACCGAGGCCCTTGACTTACTCTAGTCTGAA	9385		
Qy	1177	AGAGGGGAACCTGTTTATTTTAGAGAAAGACGCTGTATTATGTTATCAATCCAGA	1236		
Db	9384	AAAGGAGGACTCTGTATATTCTTAAATGAAAGAGTGTGTTTATACCTAAATCAATCTGGC	9325		
Qy	1237	ATTGTCACTCAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGC- --AGAGGAG	1293		
Db	9324	CTGGTGATGACAAACATAAGAAACTCAAGATAGGGTCCAAAACTTGCCAACCAAGCA	9265		
Qy	1294	CTTCAAAACACCGAAGCTGGGCGTCTCAGCCAAATGGAATGCCCTGGGTTCTCCCTTC	1353		
Db	9264	AGTAATTTATGCTGAACGCCCTTGACACTCTCTAAITTGGATGTACTGGGTCCTCCCAATT	9205		
Qy	1354	TTAGGACCTCTAGCAGCTCTAAATATTGTTACTCCTCTTTGGACCCCTGTATCTTTTAACCTC	1413		
Db	9204	CTTAGTCCTTTAATACCTATTTTTTCTCCCTTCTTTTATTCGACACTGTATCTGCCGTTTA	9145		
Qy	1414	CTTGTTAAGTTTGTCTCTCCAGAAATGGAAGCTGTAAAGCTACAGATGGTCTTACAAATG	1473		
Db	9144	GTCTTCTCAATTCTCCAAAACCGTATCCAGGCCAATCACCAAATCAATCTATACAAACAAATG	9085		

RESULT 28
US-09-949-016-184497/C
; Sequence 184497, Application US/09949016
; Patent No. 6812139
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION
; AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

```

, PRIOR APPLICATION NUMBER: 60/241,755
, PRIOR FILING DATE: 2000-10-20
, PRIOR APPLICATION NUMBER: 60/237,768
, PRIOR FILING DATE: 2000-10-03
, PRIOR APPLICATION NUMBER: 60/231,498
, PRIOR FILING DATE: 2000-09-08
, NUMBER OF SEQ ID NOS: 207012
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 184497
, LENGTH: 601
, TYPE: DNA
, ORGANISM: Human
US-09-949-016-184497

```

Query Match	7.9%; Score 116.6; DB 4; Length 601;
Best Local Similarity	58.5%; Pred. No. 3.5e-27;
Matches	257; Conservative 1; Mismatches 175; Indels 6; Gaps 3;
Qy	524 TCCATGAGGTCTCAGCCCAAAACCCCTACTAACTGTTGGATGTGCTCCCTCGCAGCTTCA 583
Db	599 TACAGAAGCCTCTCCTGGTAATCCAACCAACTGCTGGATGTCTCCCTTGCA-TTTT 541
Qy	584 GGCATACATTTCAATCCCTGTTCTCTGGAAACAATGGAAACAATTCAGCACAGAAATAACA 643
Db	540 CAATGTATGTCCCAAGTCCCTGTCCCTGGACAGTGGAACTTATCACCCTCAGTACTAAACA 481
Qy	644 CCATCTCCGTTTATAGTAGGACCTCTGTGTTCCAATCTGGAATAAACCCATACCTCAAAAC 703
Db	480 TCA---CCAATCAATCAGTCCCATAGTCACCAATTTACGAGCCACACAGGCTCAATC 424
Qy	704 TCACCTGTGTAAAAATTTAGCAATACTATAGACACAACAGCTCCCAATGCATCAGGTGGG 763
Db	423 TCAGTGCATAAACTTCAGCATGGCTGTCAATAAGAAACACCTTCGTGATGCAGTCCCTGTA 364
Qy	764 TACACCTCCACACGAATAGTCGCTACCTACCTCAGGAATATTTTTTGTCTGTGTAACCT 823
Db	363 TATCAGTAACCCAGGTTTCACCTAACTTCAGGCA--TTTTTTTTTCACTGUGAATTACA 306
Qy	824 CAGCTATCATTTGTTGAATGGCTCTTCAGAATCATGTGCTTCCTCTCATTTTAGTGC 883
Db	305 CAGSSTATGTATGCCTAAACAGCACTCTGAAAGAACTATGCTTCACTCCTCTTCAGCAC 246
Qy	884 CCCTATGACCATCTACACTGAACAAGATTATACAATCATGTCTGCTACCTAAGCCCCACA 943
Db	245 CTCCCATGTCCATATATACTGAACGAGAGTTACAAGTCTCCTTATACCCCAATCTGCC 186
Qy	944 ACAAAAGAGTACCCATTCT 962
Db	185 ACATGAGCCTTTTATTGT 167

```

RESULT 29
US-09-573-080A-30
; Sequence 30, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC P
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 7713
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(7713)
; OTHER INFORMATION: herbv
; PUBLICATION INFORMATION:

```

```

; PUBLICATION INFORMATION:
;
; AUTHORS: Jurka, J; Walichewicz, J; Milosavljevic, A
;
; TITLE: Prototypic sequences for human repetitive DNA
;
; JOURNAL: Journal of Molecular Evolution
;
; VOLUME: 35
;
; ISSUE: 4
;
; PAGES: 286-291
;
; DATE: 1992-10-
;
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
;
; DATABASE ENTRY DATE:
;
; DATABASE ENTRY DATE: 1996-01-26
;
; US-09-573-080A-30

```

Query Match	7.8%;	Score 115.2;	DB 4;	Length 7713;
Best Local Similarity	54.0%;	Pred. No. 5e-26;		
Matches 259;	Conservative	0;	Mismatches 218;	Indels 3; Gaps 1;
Qy	997	CTAGGTACTGGCAATTGGCAGTATCAACAACCTCTACTCAGTTCCTTACTACAACTATCTCAA	1056	
Db	7018	CTCGGTACTGGAATAGCAGTCATTTCAACCTCTGTACAGACCTTCCTGAGCCTGTCTAAT	7077	
Qy	1057	GAATATAAATGGTGACATGGAAACAGGTCACTGACTCCCTGGTACCTTGCNAGATCAACTT	1116	
Db	7078	GACTTCTCTGCTAGCATCAAGATGTGCCAACAATTTTATCAGTCTCTCCAGGCTAAAGTT	7137	
Qy	1117	AACCTCCCTAGCAGCAGTAGTCCTTTCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAA	1176	
Db	7138	GACTCTTTAGCTGCAGTGTCTCTCAAACCGCTGAGGCCCTTGACTTACTCTAGTCTGCTGA	7197	
Qy	1177	AGAGGGGGAACCTGTTTATTTTTTAGGAGAAGACGCTGTTATTTATGTTAATCAATCCAGA	1236	
Db	7198	AAAGGAGGACTGTGTATATCTTTAAATGAAGACTGTGTGTTTTTACCTAAACAACCTGGC	7257	
Qy	1237	ATTGTCACCTCAGAAAGCTTAAGAAATTCGAGATCGAATACAATGTAGAGC--AGAGGAG	1293	
Db	7258	CTGGTGTATGACAACAATAAAAAAAGCTCAAGGAAGAAGAGACCAAAAACCTTGCCAACCAAGCA	7317	
Qy	1294	CTTTCAAAACACCGAACGCTGGGGCCCTCCTCAGCCAAATGGATGCCCTGGGTCTCCCGCTTC	1353	
Db	7318	AGTAATATGCTGAATCCCTCTGGGCACCTCTTAATTGCATGTCTCTGGGTCTCTCCCAATT	7377	
Qy	1354	TTAGGACCTCTAGCAGCTCTAATATTTGTTACTCCTCTTTGGACCGCTGTATCTTTAACCTC	1413	
Db	7378	CTTAGTCCCTTTAATACCCATTTTTTCTCCCTCTTTTATTCGGACCTGTATCTCTCCCATTTA	7437	
Qy	1414	CTTTGTTAAGTTTGCTCTCTCCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTTACAAATG	1473	
Db	7438	CTTTCTCAATTCACCAACCTCTATCCAGGCATACCAATCATTTTATACGACAAATG	7497	

```

1  RESULT 30
2  US-09-573-080A-44
3  Sequence 44, Application US/09573080A
4  Patent No. 6828097
5  GENERAL INFORMATION:
6  APPLICANT: JOGAN, KNOLL
7  APPLICANT: JOGAN, PETER
8  TITLE OF INVENTION: SINGLE COPY GEN
9  FILE REFERENCE: 30307
10 CURRENT APPLICATION NUMBER: US/09/57
11 CURRENT FILING DATE: 2000-05-16
12 NUMBER OF SEQ ID NOS: 479
13 SOFTWARE: PatentIn version 3.0
14 SEQ ID NO 44
15 LENGTH: 7392
16 TYPE: DNA
17 ORGANISM: Homo sapiens
18 FEATURE:
19 NAME/KEY: repeat_region
20 LOCATION: (1)..(7392)
21 OTHER INFORMATION: herv
22 PUBLICATION INFORMATION:
23 PUBLICATION INFORMATION:

```

; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUES: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; DATABASE ENTRY DATE: 1996-01-26
; US-09-573-080A-44

Query Match 7.8%; Score 115; DB 4; Length 7392;
Best Local Similarity 56.9%; Pred. No. 5.6e-26;
Matches 211; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 1045 AAATCTCTCAAGAAATAAATGGTGACATGGAACAGGTCTACTGCTCCCTGGTGCACCTTG 1104
DB AAATTAATCTAATCAGCTAATTTCTGATGTACAAATCTTATCTAGCACCATACAAAGATCTG 6963

QY 1105 CAAGATCAACTTAATCCCTTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTG 1164
DB CAAGATCAAGTACTGACTCATTTAGCCGAGTGGTTCTCCAGAACAGAGGGGGCTAGATCTTA 7023

QY 1165 CTAACGCCAAAAGAGGGGAACCTGTTATTTTAGGAGAAGAACGCTGTTATATGTT 1224
DB CTTACAGCAGAACAGGAGGAGATCTGTTTAGCCCTGCAAGAAATGCTGTTTATGTT 7083

QY 1225 AATCAATCCAGAATTTGCTACTGAGAAGTTAAAGAAATTCAGATCGAATACAAATGTAGA 1284
DB AACAGTCAAGGATTTGAGAGACAAATAAATAAACCTTACAGAGAACTAGAGAAAGAGCT 7143

QY 1285 GCAGAGAGCTTCAAAACACCGAAGCGTGGGGCTCTCAGCCAAATGGATGCCCTGGTT 1344
DB AGAAAAGATCTAGCTTCCCAACCCACTTTGGACTGGGCTTCAAGGGCTCTCCCTTACCTC 7203

QY 1345 CTCCTCTCTTAGGACCTCTAGCAGCTCTAATATTTGTTACTCTCTTTGACCTCTGTATC 1404
DB CTGCTCTCTTTGCGCTCTCTACTTACCCTCTCTGCTCTTCTTACCATTGGGCGTGCATT 7263

QY 1405 TTTAACTCTCT 1415
DB TTTAACTCTCT 7274

RESULT 31
US-09-949-016-184495/c
; Sequence 184495, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184495
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-184495

Query Match 7.8%; Score 114.8; DB 4; Length 601;
Best Local Similarity 61.8%; Pred. No. 1.4e-26;
Matches 215; Conservative 1; Mismatches 128; Indels 4; Gaps 2;

QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
DB AGTACTCCAGTCCATACAGAAATTTAGACCTTTCCAGGCTACAGGAARCCCTTAACTCT 290

QY 481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCCTCAGCTCGGCTCCATGAGTCTCAGCC 540
DB CATTTCTCTCTCGGAGCCTCTTTAAACACCCCTTTACAGGAATACAGAAGCCTCTCCT 230

QY 541 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
DB GGTAAATCCAAACCACTGCTGGATGTCCTCCCTTGCA-TTTTCATGATGTCCTCCAGTC 171

QY 601 CTGTTTCTCTGAACAATGGAACTTTCAGCACAGAAATAAACAACCACTTCCGTTTTAGTA 660
DB CCTGCTCCCTGGACAGTGAATTTATCCACCCAGTACTAAACATCA---CCAATCAATC 114

QY 661 GGACCTCTGTTTCCATCTGGAAATAACCATACCTCAAACTCAGCTGTGTAAAATTT 720
DB AGTCCCATAGTCAACCAATTTACCAGCCACACAGGCTCAAACTCTCAGTGCATATAAACTTC 54

QY 721 AGCAATPACTATAGACAACAACAGCTCCCAATGCATCAGGTGGGTAACA 768
DB AGCATGGCTGTCATAGAACACCTTCTGATGTCACTCTGTATATCA 6

RESULT 32
US-09-078-294-4
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
; US-09-078-294-4

Query Match 7.8%; Score 114.8; DB 3; Length 80246;
Best Local Similarity 55.7%; Pred. No. 2.9e-25;
Matches 263; Conservative 0; Mismatches 202; Indels 7; Gaps 2;

QY 1010 TTGGCAGTATCAACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGGTG 1069
DB TTAACAGCTPAACATGATTAATTTCTACTACCTCAGCTCTCTCAAGAGATTTCTCAGACA 58568

QY 1070 ACATGGAACAGGTCACTGACTCCCTGGTCACTTGGCAAGTCACTTAATCCCTAGCAG 1129
DB GTTTGCAAAAAGAAACGAAATCTGCTTACTCTACAATCCCAATAGACTCTTTTGGCAG 58528

QY 1130 CAGTAGTCTTCAAAATCGAAGAGCTTAGACTTGTAAAGCCCAAGAGGGGGAACCT 1189
DB CAGTACTCTCCAAACCCGCTGAGGCTGAGCTCTTCTTACTGTGAGAAAGAGAGATTTCT 58688

QY 1190 GTTTATTTTAGGAGAAGAACGCTGTTATTTATGTTATCAATCAATCAGAAATTTGTCTGAGA 1249
DB GCAC-TTCTTAGGGTAGAGTGTGTTTTTATTAACCACTCAGGAGTAATATGAGATA 58747

QY 1250 AAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGGAGCTTTCAAAACACCAAC 1309
DB CCACCCAGTCTTTTACAGGAAAGGCTTCTGAAATCAGACAAATGCTTTCAAACTCTTATA 58807

QY 1310 GCTGGGGCCCTCTCAGCCAAATGGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAG 1369
DB CCAA-----CCTTGAGTTGGGCGACATGGGCTTCTCCCTCTTCTAGGCTCTGTGACAG 58861

QY 1370 CTCTAATATGTTACTCTCTTGGACCGCTGTATCTTTAACTCCTCTTGTAAAGTTTGTCT 1429
 Db 58862 CCACTCTGCTAATAGTCGCAATTTGGCCCTGTATTTTAACTCTTGTGCAAAATTTGTTT 58921
 QY 1430 CTTCCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481
 Db 58922 CCTCTAGGATCGAGGCCATCAAGCTACAGATGATCTTCAAAATGTAACCCCA 58973

RESULT 33

US-09-011-745-1
 ; Sequence 1, Application US/09011745
 ; Patent No. 6165715
 ; GENERAL INFORMATION:
 ; APPLICANT: Collins, Mary KL
 ; APPLICANT: Weiss, Robin A
 ; APPLICANT: Takeuchi, Yasuhiro
 ; TITLE OF INVENTION: Expression systems
 ; FILE REFERENCE: 09/011,745
 ; CURRENT APPLICATION NUMBER: US/09/011,745
 ; EARLIER FILING DATE: 1998-06-22
 ; EARLIER APPLICATION NUMBER: PCT/GB96/02061
 ; EARLIER FILING DATE: 1996-08-23
 ; EARLIER APPLICATION NUMBER: GB9517263.1
 ; EARLIER FILING DATE: 1995-08-23
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2518
 ; TYPE: DNA
 ; ORGANISM: RD114
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)
 ; OTHER INFORMATION: n is any nucleotide
 ; US-09-011-745-1

Query Match 7.7%; Score 114.2; DB 3; Length 2518;
 Best Local Similarity 55.0%; Pred. No. 5.2e-26;
 Matches 224; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1045 AAATCTATCTCAAGAAATAAATGTTGACATGGAACAGGTCACCTGCTCCCTGTCACCTTG 1104
 Db 1834 AAATATATCCCATCAGTTAATATCTGATGTCCAAAGCTTTATCCGGTACCATACAAGATTTA 1893
 QY 1105 CAAGATCAACTTAACTCCCTAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTG 1164
 Db 1894 CAAGACGAGTAGACTGTTAGCTGAAGTAGTTCTCCAAATAGGAGGGGACTGGACCTA 1953
 QY 1165 CTAACCGCCAAAAGAGGGGAACCTGTTTATTTTATAGGAGAAACGCTGTTTATATGTT 1224
 Db 1954 CTAAGGSCAGACAAGGAGGAAATTTGTTTAGCCTTACAGAAAAATGCTGTTTATGCT 2013
 QY 1225 AATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATAGA 1284
 Db 2014 AACAAGTCAGGAATTTGAGAAAACAAATTAAGAACCCCTACAAGAAATTTACAAAACCC 2073
 QY 1285 GCAGAGAGCTTCAAAACACCGACGCTGGGGCTCTCTCAGCCCAATGGATGCCCTGGGTT 1344
 Db 2074 AGGGAAGCTGGCAACCAACCTCTCTGACCGGGCTGCAGGGCTTCTTCGGTACCTC 2133
 QY 1345 CTCCTCTCTTAGACCTCTAGACGCTCTAATATGTTTACTCTCTTTGGACCTGTATC 1404
 Db 2134 CTACCTCTCTGGGACCCCTACTCACCTCTCTACTCATACTAACCATTGGGCCATCGGTT 2193
 QY 1405 TTTAACTCTCTTGTAAAGTTGTTCTCTCCAGAAATTCAGCTGATAA 1451
 Db 2194 TTACGTGCCTCATGGCCCTTCATTAATGATAGACTTAATGTTGACA 2240

RESULT 34

US-09-011-745-8
 ; Sequence 8, Application US/09011745
 ; Patent No. 6165715
 ; GENERAL INFORMATION:
 ; APPLICANT: Collins, Mary KL
 ; APPLICANT: Weiss, Robin A
 ; APPLICANT: Takeuchi, Yasuhiro
 ; APPLICANT: Cosset, Francois-Loic
 ; TITLE OF INVENTION: Expression systems
 ; FILE REFERENCE: 09/011,745
 ; CURRENT APPLICATION NUMBER: US/09/011,745
 ; CURRENT FILING DATE: 1998-06-22
 ; EARLIER APPLICATION NUMBER: PCT/GB96/02061
 ; EARLIER FILING DATE: 1996-08-23
 ; EARLIER APPLICATION NUMBER: GB9517263.1
 ; EARLIER FILING DATE: 1995-08-23
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 5865
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Portion of
 ; OTHER INFORMATION: construct
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3611)
 ; OTHER INFORMATION: n is any nucleotide
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3612)
 ; OTHER INFORMATION: n is any nucleotide
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3613)
 ; OTHER INFORMATION: n is any nucleotide
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3614)
 ; OTHER INFORMATION: n is any nucleotide
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3799)
 ; OTHER INFORMATION: n is any nucleotide
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3800)
 ; OTHER INFORMATION: n is any nucleotide
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3801)
 ; OTHER INFORMATION: n is any nucleotide
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (3802)
 ; OTHER INFORMATION: n is any nucleotide
 ; US-09-011-745-8

Query Match 7.7%; Score 114.2; DB 3; Length 5865;
 Best Local Similarity 55.0%; Pred. No. 8.8e-26;
 Matches 224; Conservative 0; Mismatches 183; Indels 0; Gaps 0;
 QY 1045 AAATCTATCTCAAGAAATAAATGTTGACATGGAACAGGTCACCTGCTCCCTGTCACCTTG 1104
 Db 2692 AAATATATCCCATCAGTTAATATCTGATGTCCAAAGCTTTATCCGGTACCATACAAGATTTA 2751
 QY 1105 CAAGATCAACTTAACTCCCTAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTG 1164
 Db 2752 CAAGACGAGTAGACTCTGTTAGCTGAAGTAGTTCTTCCAAAATAGAGGGGACTGGACCTA 2811
 QY 1165 CTAACCGCCAAAAGAGGGGAACTGTTTATTTTAGGAGAAACGCTGTTTATTATGTT 1224

Db 2812 CTAACGCGACAAAGGAGGAATTTGTTTAGCCTTACAGAAAATGCTGTTTATGCT 2871
QY 1225 AATCAATCCAGANTTGTCTAGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGAGA 1284
Db 2872 AACAAGTCAGGAATTTGAGAAAACAAATTAAGAACCTTACAGAGAAATTTACAAAACGC 2931
QY 1285 GCAGAGAGCTTCAAAAACACCGAACGCTGGGGCTCTCTCAGCCAATGGATGGCTGGGTT 1344
Db 2932 AGGGAAGCTGGCAACCAACCTCTCTGGACGGGCTGCAGGGCTTCTTCGCTACCTC 2991
QY 1345 CTCCTCTCTTAGACCTCTAGCAGCTCTAATATTTGTTACTCTCTTTGGACCTGTATC 1404
Db 2992 CTACCTCTCTGGACCTCTACTCACCTCTCTACTCATACTAAACCAATGGGCGCATGGTT 3051
QY 1405 TTTAACTCTCTTGTAAATTTGTTCTCTCCAGAAATGAAGCTGAAA 1451
Db 3052 TTCAGTCGCTCATGGCCTTCATTAATGATAGACTTAATGTTGTACA 3098

RESULT 35

US-09-573-080A-28
; Sequence 28, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATION
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573.080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28
; LENGTH: 6529
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)...(6529)
; OTHER INFORMATION: herfth2
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-28

Query Match 7.6%; Score 112.6; DB 4; Length 6529;
Best Local Similarity 54.3%; Pred. No. 3.1e-25;
Matches 274; Conservative 0; Mismatches 224; Indels 7; Gaps 2;
QY 929 TACTTAGCCCCCAACAAAGAGTACCCTATTTCTCTTTTGTATCAGAGCAGGATGC 988
Db 5839 TCCCTATCTGACACGCSATCTATACAACTCATACCTCTGTAGTAACCCCTCAGAATAA 5898
QY 989 TAGGAGACTAGTACTGGCAATGGCAGTATCACAACTCTACTAGTTCTACTACAAC 1048
Db 5899 CTACAGGAGTTGGAACTGGGATTTGGGATTAACCACTCTCTTCCCTATTACCAATCCC 5958
QY 1049 TATCTCAGAAATAAATGGTCAGATGGAACAGGTCACTGACTCTCTGGTCACTTGCAG 1108
Db 5959 TCTCAAAGACTTCAAGATAGCTTGGAGAGATAGCCAAATCCATCAACTCTCCAT 6018
QY 1109 ATCAACTTAACTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGTCAA 1168
Db 6019 CACAAATAAATCT-TAGCAGCAGTAGTCTTCAAAACCGCAGAGGCTTAGACCTGTTC 6077

QY 1169 CGCCCAAAAGAGGGGAAACCTGTTTATTTTATAGGAGAAAGCGCTGTTTATTTATGTTAATC 1228
Db 6078 CAGCGAAAAAGAGAACTCTGCTTTTCTAGATGAACAGATATGCTTTTATCTTAACC 6137
QY 1229 AATCCAGAAATTCCTACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAG 1288
Db 6138 AATCTGGCCTGTATACAGGATCTGTAAAGAAAGACTAAAGACCGAGCACAAAATTAAG 6197
QY 1289 AGGAGCTTCAAAAACACCGAACGCTGGGGCTC-----CTCAGCCAATGGATGCCCTCGG 1342
Db 6198 AAAACGTCTCTGATGGCCAGCGTGGCCCTCTCTGGTCTTTTAGTACCTGCTTTCCATGGC 6257
QY 1343 TTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTTGTTATCTCTCTTTTGGACCTGTGA 1402
Db 6258 TACTGCCCTCTCTAGGCCCTGCGGTAACCAATTTTCTCTTTCTAGCATTTGGCCCTGTGC 6317
QY 1403 TCTTTAACTCTCTTGTAAATTTGT 1427
Db 6318 TCTTACATCTCTTACCAGTTTTT 6342

RESULT 36

US-09-555-352-1
; Sequence 1, Application US/09555352
; Patent No. 6544779
; GENERAL INFORMATION:
; APPLICANT: Cichutek, Klaus
; APPLICANT: Mergel-Millitzer, Heike
; TITLE OF INVENTION: PSEUDO-TYPE RETROVIRAL VECTORS WITH
; TITLE OF INVENTION: MODIFIABLE SURFACE CAPSID PROTEINS
; FILE REFERENCE: 11692-005001
; CURRENT APPLICATION NUMBER: US/09/555,352
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/DE98/03542
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: DE 197 52 855.4
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4776
; TYPE: DNA
; ORGANISM: Murine leukemia virus
US-09-555-352-1

Query Match 7.6%; Score 112.2; DB 4; Length 4776;
Best Local Similarity 51.5%; Pred. No. 3.5e-25;
Matches 258; Conservative 0; Mismatches 243; Indels 0; Gaps 0;
QY 951 AGTACCCTATCTTCTCTTTTCTTATCAGAGCAGAGTGCTAGGCAGACTAGGTACTGGCAT 1010
Db 1216 AGTCCAGTTTATCCCTCTGCTTGTGGGTCTAGGGATTCAGGGGCTACACTTGTCTGGTGG 1275
QY 1011 TGGCAGTATCACAAACCTCTACTCAGTTCTACTACAACTACTCAAGAAATAAATGTGTA 1070
Db 1276 AACGGGCTTGGGGTCTCCGCTTACACTTATCAAGCTCTCTAATCAATGTTGAAGA 1335
QY 1071 CATGGAAACAGTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAATCCCTTAGCAGC 1130
Db 1336 TGTCAGGCTCTTTTCAAGGACCAATCAATGACCTACAGGACAGATTGACTCCCTGGCTGA 1395
QY 1131 AGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCAAAAGAGGGGAAACCTG 1190
Db 1396 GGTTGTCTTACAAATAGAGAGGGTTAGACCTATTGACTTGCAGAACAGAGGAGGAATG 1455
QY 1191 TTTATTTTATAGGAGAAAGACGCTGTTTATTTATGTTAATCAATCCAGAAATTTGCTACTGAGAA 1250
Db 1456 TCTCGCACTCCAGGAGAGTGTGTTTACGCTTAAACAGTCCGGTATCGTACGTGACAA 1515
QY 1251 AGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCGAAGC 1310
Db 1516 GATCCGAAAACCTCCAGAGGACCTTATCGAGAGAAAACGTCGACTGTACGACAAACCCCT 1575

QY 1311 CTGGGCGCTCTCAGCCAAATGATGCCCTGGGTTCTCCCGTTCTTAGACCTTAGCAGC 1370
Db 1576 GTGGAGCGCTTGAACGGCTTCCCTTCCATATTTGCTACCGCTTGTAGGCCCGCTGTTGG 1635
QY 1371 TCTAATATTGTTACTCCTCTTTGGACCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTC 1430
Db 1636 GCTCATATTGTTCTTGAACCTCGGCCGCGTGCATTATGAAGACCTTGACTCGCATTTATACA 1695
QY 1431 TTCCAGAATTGAAGCTGTAAA 1451
Db 1696 TGACAAATTCAGGCAGTAAA 1716

RESULT 37

US-09-620-312D-171
; Sequence 171, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 171
; LENGTH: 2809
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1936)..(2364)
US-09-620-312D-171

Query Match 7.2%; Score 106.4; DB 4; Length 2809;
Best Local Similarity 55.0%; Pred. No. 1.9e-23;
Matches 231; Conservative 0; Mismatches 186; Indels 3; Gaps 1;
QY 996 ACTAGGTACTGGCAGTATCAACCTCTACTAGTTCTACTACAAACTATCTCA 1055
Db 1851 ACTTGGCATTGGATAGCAGGCAATACCCTCAATCACTCATACCAACACTATTCAC 1910
QY 1056 AGAAATAAATGGTGACATGGAAACAGGTCACTGACTCCCTGGTCACTTGGCAAGATCAACT 1115
Db 1911 AACCTTTCTAACACCGTAGAATATGCACACTTCCATTACCAGTCTCCAAACGCAAT 1970
QY 1116 TAACCTCCCTAGCAGCTAGTCTTCAAAATCGAAGCTTTAGACTTGTACCTGCTTAACGCCAA 1175
Db 1971 AGACTTCTCTGGGAGTCACTCTTCAAACTGGAGATCCTGGACCTCTTAACCACTGA 2030
QY 1176 AAGAGGGGGAACCTGTTTATTTTAGGAGAAGAACGCTGTTTATTTATGTTAAATCCAG 1235

Db 2031 GAAAGGGGTACCTGCATATACCTCCAGGAAGAAATGCTGTTTCTGTTTAATGAATCTGG 2090
QY 1236 AATTGTCACGTAGAAAGTTAAAGAAATTCGAGATCG--AATACAATGTAGAGCAGAGGA 1292
Db 2091 CATTTGTTTCATATCGCAGTTTCGTAGGCTTCATGACAGGGCTGCAGAGCTTAGACATCAAGT 2150
QY 1293 GCTTCAAAACACCGAACGCTGGGGCTCTCAGCCAATGGATGCCCTGGGTTCTTCCCGCTT 1352
Db 2151 CGCTGACTCTGGTGGCAAGATCATCCCTTCTAAGATGATACCTTGGGTTGGCCCGCTT 2210
QY 1353 CTTAGGACCTCTAGCAGCTCTAATATTGTTACTCTCTTTTGGACCTGTATCTTTAACTT 1412
Db 2211 CCTAGGACCCCTGATCTTCTCTCTGTTACTAATGATTGGGCCATGTCATATTTAACTT 2270

RESULT 38

US-09-904-615-12
; Sequence 12, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-615-12

Query Match 7.1%; Score 104.8; DB 4; Length 2342;
Best Local Similarity 54.8%; Pred. No. 5.7e-23;
Matches 230; Conservative 0; Mismatches 187; Indels 3; Gaps 1;
QY 996 ACTAGGTACTGGCAGTATCAAACTCTACTCAGTTCTTACTACAAACTATCTCA 1055
Db 1357 ACTTGGCAGTGGATAGCAGGCATAAACCTCTCAATCACTCATACCAACACTATTCCAC 1416
QY 1056 AGAAATAAATGGTGACATGGAAACAGGTCACTGACTCCCTGGTCACTTGGCAAGATCAACT 1115
Db 1417 AACCTTTTAAACACCGTAGAAGATATGCACACTTCCATTACCAGTCTCCAAACGCAAT 1476
QY 1116 TAACCTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGACTTTAGACTTTGCTAAACGCCAA 1175
Db 1477 AGACTTCTCTGGGAGTCACTCTTCAAACTGGAGATCTCTGGACCTCTTAACCACTGA 1536
QY 1176 AAGAGGGGGAACCTGTTTATTTTAGGAGAAGAACGCTGTTTATTTATGTTAAATCCAG 1235
Db 1537 GAAAGGGGTACTCTGCATATACCTCCAGGAAGAAATGCTGTTTCTGTTTAATGAATCTGG 1596
QY 1236 AATTGTCACGTAGAAAGTTAAAGAAATTCGAGATCG--AATACAATGTAGAGCAGAGGA 1292
Db 1597 CATTTGTTTCATATCGCAGTTTCGTAGGCTTTCATGACAGGGCTGCAGAGCTTTGACATCAAGT 1656
QY 1293 GCTTCAAAACACCGAACGCTGGGGCTCTCAGCCAATGGATGCCCTGGGTTCTCCCGCTT 1352
Db 1657 CGCTGACTCTCTGGTGGCAAGGATCATCCCTTCTAAGATGGATACCTCTGGGTTGGCCCGCTT 1716
QY 1353 CTTAGGACCTCTAGCAGCTCTAATATTGTTACTCTCTTTTGGACCTGTATCTTTAACTT 1412
Db 1717 CCTAGGACCCCTGATCTTCTCTCTGTTACTAATGATTGGGCCATGTCATATTTAACTT 1776

RESULT 39
 US-09-949-016-12735/c
 ; Sequence 12735, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12735
 ; LENGTH: 80355
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-12735

Query Match 7.0%; Score 103.4; DB 4; Length 80355;
 Best Local Similarity 53.1%; Pred. No. 1.5e-21;
 Matches 246; Conservative 0; Mismatches 211; Indels 6; Gaps 1;
 QY 995 GACTAGTACTGGCATTGGCAGTATCAACCTCTACTCAGTTCTACTCAAACTATCTC 1054
 Db 59660 GAGGTGAACTGAGACTGCGGGATTACCACTCTGTTCTTATTAACCAATCCCTCTCCA 59601
 QY 1055 AAGAAATAAATGGTGAATGGAACAGTCACTGACTCCCTGGTCAAGATCAAC 1114
 Db 59600 AAGACCTCACAGATAGCTTGGAAAGATAGCCAAATCCATCAACTCTCCAATCAAAA 59541
 QY 1115 TTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCA 1174
 Db 59540 TAGATTCTTTAGCAGCAGCGCTTCTTCAAAATCACAGAGCGGTAGACCGACTCACAGCCA 59481
 QY 1175 AAAGAGGGGAACTGTTTATTTTATAGGAGAAAGACGCTGTTTATTTATGTTAAATCAATCCA 1234
 Db 59480 AAAGAGGAGACTCTGCTCTTTCTAGATGAACAGTGTGCTTTTATCGTAACCAATCTG 59421
 QY 1235 GAATTGCTACTGAGAAAGTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGGAGC 1294
 Db 59420 GCTTAGTACGAGATTTTGTAAATAAACAATAAGGACCGAGCAAAAAATTAAGAAAAATG 59361
 QY 1295 TTC-----AAACACCGAAGCTGGGGCTCTCAGCCAAATGGATGCCCTGGGTTCTCC 1348
 Db 59360 TCCCCCGATGGCCAGCGTGGCCCTCTGTCCTTTAGTACCTGGTTTCCATGGCTAATGC 59301
 QY 1349 CCTTCTTAGGACCTCTAGCAGCTCTAATATGTTTACTCCTCTTTGGACCCCTGTATCTTTA 1408
 Db 59300 CCTTCTTAGGCCAGCATAACCAATCTTCTTTTCTAGCATTGAGCCCTGTCTCTCTAC 59241
 QY 1409 ACCTCCTTGTAGTTGTCTCTTCCAGAAATTTGAAGCTGTAAA 1451
 Db 59240 GCCTTCTCACTCAGTTTTTTACAGGACCGTATCGGAGCCTTCAA 59198

RESULT 40
 US-09-949-016-13572/c
 ; Sequence 13572, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13572
 ; LENGTH: 80357
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-13572

Query Match 7.0%; Score 103.4; DB 4; Length 80357;
 Best Local Similarity 53.1%; Pred. No. 1.5e-21;
 Matches 246; Conservative 0; Mismatches 211; Indels 6; Gaps 1;
 QY 995 GACTAGTACTGGCATTGGCAGTATCAACCTCTACTCAGTTCTACTCAAACTATCTC 1054
 Db 59660 GAGGTGAACTGAGACTGCGGGATTAAACCACTCTGTTCTTATTAACCAATCCCTCTCCA 59601
 QY 1055 AAGAAATAAATGGTGAATGGAACAGTCACTGACTCCCTGGTCAAGATCAAC 1114
 Db 59600 AAGACCTCACAGATAGCTTGGAAAGATAGCCAAATCCATCAACTCTCCAATCAAAA 59541
 QY 1115 TTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCA 1174
 Db 59540 TAGATTCTTTAGCAGCAGCGCTTCTTCAAAATCACAGAGCGGTAGACCGACTCACAGCCA 59481
 QY 1175 AAAGAGGGGAACTGTTTATTTTATAGGAGAAAGACGCTGTTTATTTATGTTAAATCAATCCA 1234
 Db 59480 AAAGAGGAGACTCTGCTCTTTCTAGATGAACAGTGTGCTTTTATCGTAACCAATCTG 59421
 QY 1235 GAATTGCTACTGAGAAAGTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGGAGC 1294
 Db 59420 GCTTAGTACGAGATTTTGTAAATAAACAATAAGGACCGAGCAAAAAATTAAGAAAAATG 59361
 QY 1295 TTC-----AAACACCGAAGCTGGGGCTCTCAGCCAAATGGATGCCCTGGGTTCTCC 1348
 Db 59360 TCCCCCGATGGCCAGCGTGGCCCTCTGTCCTTTAGTACCTGGTTTCCATGGCTAATGC 59301
 QY 1349 CCTTCTTAGGACCTCTAGCAGCTCTAATATGTTTACTCCTCTTTGGACCCCTGTATCTTTA 1408
 Db 59300 CCTTCTTAGGCCAGCATAACCAATCTTCTTTTCTAGCATTGAGCCCTGTCTCTCTAC 59241
 QY 1409 ACCTCCTTGTAGTTGTCTCTTCCAGAAATTTGAAGCTGTAAA 1451
 Db 59240 GCCTTCTCACTCAGTTTTTTACAGGACCGTATCGGAGCCTTCAA 59198

Search completed: February 21, 2005, 15:01:25
 Job time : 267.948 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 11:52:50 ; Search time 818.526 Seconds
(without alignments)
10694.149 Million cell updates/sec

Title: US-09-319-156B-9
Perfect score: 1481
Sequence: 1 atggccctcccttatcatc.....gtctacaagtgaacccca 1481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5384158 seqs, 2955248155 residues
Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	1481	US-08-979-847-105	Sequence 105, App
2	1481	100.0	1481	US-10-114-104-105	Sequence 105, App
3	1481	100.0	2030	US-10-637-565-18	Sequence 18, Appl
4	1389.8	93.8	2074	US-10-416-642-4	Sequence 4, Appl
5	1335.4	90.2	2930	US-09-902-535-1	Sequence 1, Appl
6	1335.4	90.2	2946	US-10-114-893-134	Sequence 134, App
7	1335.4	90.2	2946	US-10-016-249-3	Sequence 3, Appl
8	1333.8	90.1	1617	US-10-133-036-3	Sequence 3, Appl
9	1333.8	90.1	8523	US-09-854-867-21	Sequence 21, Appl
10	1333.8	90.1	56093	US-09-873-367C-81	Sequence 81, Appl
11	1324.2	89.4	7582	US-10-632-793-30	Sequence 30, Appl

12	1308.2	88.3	2782	17	US-10-133-036-1	Sequence 1, Appl
13	1306.6	88.2	2782	17	US-10-632-793-26	Sequence 26, Appl
14	1277.4	86.3	161334	13	US-10-087-192-730	Sequence 730, App
15	1131.6	76.4	1894	9	US-09-864-761-4444	Sequence 4444, Ap
16	1116.4	75.4	1948	17	US-10-632-793-24	Sequence 24, Appl
17	716.6	48.4	792	9	US-09-864-761-21192	Sequence 21192, A
18	591.8	40.0	1684	17	US-10-363-616-228	Sequence 228, App
19	582.8	39.4	591	17	US-10-632-793-19	Sequence 19, Appl
20	516.6	34.9	1136	17	US-10-632-793-25	Sequence 25, Appl
21	513.6	34.7	2006	17	US-10-632-793-23	Sequence 23, Appl
22	377.8	25.5	570	13	US-10-027-632-322491	Sequence 322491, A
23	377.8	25.5	570	13	US-10-027-632-322492	Sequence 322492, A
24	377.8	25.5	570	17	US-10-027-632-322491	Sequence 322491, A
25	377.8	25.5	570	17	US-10-027-632-322492	Sequence 322492, A
26	364	24.6	521	16	US-10-029-386-4312	Sequence 4312, Ap
27	353	23.8	551	13	US-10-027-632-322574	Sequence 322574, A
28	353	23.8	551	13	US-10-027-632-322575	Sequence 322575, A
29	353	23.8	551	17	US-10-027-632-322574	Sequence 322574, A
30	353	23.8	551	17	US-10-027-632-322575	Sequence 322575, A
31	318.6	21.5	822900	17	US-10-292-798-1393	Sequence 1393, Ap
32	302	20.4	1329	8	US-08-979-847-108	Sequence 108, App
33	302	20.4	1329	16	US-10-114-104-108	Sequence 108, App
34	296	20.0	426	9	US-09-864-761-14030	Sequence 14030, A
35	229.8	15.5	625	13	US-10-027-632-316412	Sequence 316412, A
36	229.8	15.5	625	17	US-10-027-632-316412	Sequence 316412, A
37	229.6	15.5	619	13	US-10-027-632-86541	Sequence 86541, A
38	229.6	15.5	619	17	US-10-027-632-86541	Sequence 86541, A
39	219.4	14.8	559	9	US-09-864-761-7501	Sequence 7501, Ap
40	213	14.4	494	16	US-10-029-386-20259	Sequence 20259, A
41	194.8	13.2	520	9	US-09-864-761-7310	Sequence 7310, Ap
42	189.8	12.8	740	13	US-10-027-632-150080	Sequence 150080, A
43	189.8	12.8	740	17	US-10-027-632-150080	Sequence 150080, A
44	173.8	11.7	527	13	US-10-027-632-85218	Sequence 85218, A
45	173.8	11.7	527	17	US-10-027-632-85218	Sequence 85218, A

ALIGNMENTS

RESULT 1
US-08-979-847-105
; Sequence 105, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUREK, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC PURPOSES
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435

```

; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-979-847-105

Query Match      100.0%; Score 1481; DB 8; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTCTCTTACCCGCTTTCGCTCTCACT 60
DB 1 ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTCTCTTACCCGCTTTCGCTCTCACT 60
QY 61 GCACCCCTCCATGCTGCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 61 GCACCCCTCCATGCTGCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
QY 121 AGCGGCTTCTGGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB 121 AGCGGCTTCTGGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
QY 181 ACCTTCACCTGCCACACCCATATGCCCGCACTGCTATACTTGCACCTCTTTGCATG 240
DB 181 ACCTTCACCTGCCACACCCATATGCCCGCACTGCTATACTTGCACCTCTTTGCATG 240
QY 241 CATGCAAACTACTTATTGGACAGGAAAGATTAACTCTAGTTGCTCGAGGACTT 300
DB 241 CATGCAAACTACTTATTGGACAGGAAAGATTAACTCTAGTTGCTCGAGGACTT 300
QY 301 GGAGCCACTGTCTGTTGGACTTACTTACCCATACAGATGATGTCGATGGGGTGGAAAT 360
DB 301 GGAGCCACTGTCTGTTGGACTTACTTACCCATACAGATGATGTCGATGGGGTGGAAAT 360
QY 361 CAAGGTGAGCAAGAGAAAAACAGTAAAGGAGCAATCTCCCACTGACCCGGGACAT 420
DB 361 CAAGGTGAGCAAGAGAAAAACAGTAAAGGAGCAATCTCCCACTGACCCGGGACAT 420
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCCTCCGTACC 480
DB 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCCTCCGTACC 480
QY 481 CATACTCGGCTGTGAGCCCTATTAAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
DB 481 CATACTCGGCTGTGAGCCCTATTAAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
QY 541 CAAAACCCCTACTAACTGTTGGATGTGCCCTCCCTGCATCTCAGGCCATACATTTCAATC 600
DB 541 CAAAACCCCTACTAACTGTTGGATGTGCCCTCCCTGCATCTCAGGCCATACATTTCAATC 600
QY 601 CTGTTTCTGAAACAATGGAACAATCTCAGCACAGAAATAAACACCACTTCGCTTTTAGTA 660
DB 601 CTGTTTCTGAAACAATGGAACAATCTCAGCACAGAAATAAACACCACTTCGCTTTTAGTA 660
QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCCTGCTGTGTAATAATTT 720
DB 661 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCCTGCTGTGTAATAATTT 720
QY 721 AGCAATACTATACACACACCCAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
DB 721 AGCAATACTATACACACACCCAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
QY 781 ATAGTCTGCGCTACCCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCCTATCATTTGTTG 840
```

```

DB 781 ATAGTCTGCGCTACCCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCCTATCATTTGTTG 840
QY 841 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
DB 841 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
QY 901 ACTGAACAAGATTTATACAATCATGTCTACCTAAGCCCCACACAAACAAAGAGTACCCATT 960
DB 901 ACTGAACAAGATTTATACAATCATGTCTACCTAAGCCCCACACAAACAAAGAGTACCCATT 960
QY 961 CTTCTCTTTTGTATCAGAGCAGGAGTCTAGGAGAGACTAGGTACTTGGCATTTGGCAGTATC 1020
DB 961 CTTCTCTTTTGTATCAGAGCAGGAGTCTAGGAGAGACTAGGTACTTGGCATTTGGCAGTATC 1020
QY 1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
DB 1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
QY 1081 GTCACTGACTCCCTGGTCACTTGCAGAGTCAAACTTAACTCCCTAGCAGCAGTAGTCTCTT 1140
DB 1081 GTCACTGACTCCCTGGTCACTTGCAGAGTCAAACTTAACTCCCTAGCAGCAGTAGTCTCTT 1140
QY 1141 CAAAATCGAAGAGCTTTAGACTTGTAAACCCGCAAAAGAGGGGGAACCTGTTTATTTTA 1200
DB 1141 CAAAATCGAAGAGCTTTAGACTTGTAAACCCGCAAAAGAGGGGGAACCTGTTTATTTTA 1200
QY 1201 GGAGAAAGACGCTGTTATTTAATCAATCCAGAAATTTGTCACCTGAGAAAGTTTAAAGAA 1260
DB 1201 GGAGAAAGACGCTGTTATTTAATCAATCCAGAAATTTGTCACCTGAGAAAGTTTAAAGAA 1260
QY 1261 ATTGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAAACACCGAACGCTGGGGCCTC 1320
DB 1261 ATTGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAAACACCGAACGCTGGGGCCTC 1320
QY 1321 CTCAGCCAAATGGATGGCTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 1321 CTCAGCCAAATGGATGGCTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
QY 1381 TTACTCTCTTTGACCCCTGATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGAAAT 1440
DB 1381 TTACTCTCTTTGACCCCTGATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGAAAT 1440
QY 1441 GAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAACCCCA 1481
DB 1441 GAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAACCCCA 1481
```

RESULT 2
US-10-114-104-105
; Sequence 105, Application US/10114104
; Publication No. US20030198647A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,104
FILING DATE: 03-APR-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,847
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-10-114-104-105
Query Match 100.0%; Score 1481; DB 16; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTATCATACATTTTCTTTACTGTTCTTTACCCGCTTTCGCTCTCACT 60
Db 1 ATGGCCCTCCCTATCATACATTTTCTTTACTGTTCTTTACCCGCTTTCGCTCTCACT 60
QY 61 GCACCCCTCCATGCTCTGCTACACAGTAGTCCCTTACCAGAGTTTCTATGAAGA 120
Db 61 GCACCCCTCCATGCTCTGCTACACAGTAGTCCCTTACCAGAGTTTCTATGAAGA 120
QY 121 ACGGGGCTTCTGGAAATATTGATGCCCCCATATAGAGATTATCTAAGGGAATCC 180
Db 121 ACGGGGCTTCTGGAAATATTGATGCCCCCATATAGAGATTATCTAAGGGAATCC 180
QY 181 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATAAATCTGCACTCTTTGCATG 240
Db 181 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATAAATCTGCACTCTTTGCATG 240
QY 241 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 300
Db 241 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 300
QY 301 GGAGCCACTGCTGTGGACTTACTTCAACCATACAGATATGCTGATGGGGTGAATT 360
Db 301 GGAGCCACTGCTGTGGACTTACTTCAACCATACAGATATGCTGATGGGGTGAATT 360
QY 361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
Db 361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCAAACTACATGAACCCCTCGGTACC 480
Db 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCAAACTACATGAACCCCTCGGTACC 480
QY 481 CATACTCGCTGGTGGAGCTATTTAATACCAACCTCTCACTCGGCTCCCATGAGGTCTCAGCC 540
Db 481 CATACTCGCTGGTGGAGCTATTTAATACCAACCTCTCACTCGGCTCCCATGAGGTCTCAGCC 540
QY 541 CAAAACCCCTACTAACTGTGTGATGTCCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
Db 541 CAAAACCCCTACTAACTGTGTGATGTCCTCCCTCGCACTTCAGGCCATACATTTCAATC 600

RESULT 3
US-10-637-565-18
; Sequence 18, Application US/10637565
; Publication No. US20040043381A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: PERRON, Hervé
; APPLICANT: KOMURIAN-PRADEL, Florence
; TITLE OF INVENTION: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES
; FILE REFERENCE: 110257
; CURRENT APPLICATION NUMBER: US/10/637,565

601 CCTGTTCTGAAACAATGGAAACAACCTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
Db CCTGTTCTGAAACAATGGAAACAACCTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
QY 661 GGACCTCTGTTTCCAAATCTGGAAATAACCCATACTCAAACTCCTACCTGTGTAAATTT 720
Db 661 GGACCTCTGTTTCCAAATCTGGAAATAACCCATACTCAAACTCCTACCTGTGTAAATTT 720
QY 721 AGCAATATCTATAGACACACACAGCTCCCAATCATCAGGTGGGTAAACACTCCACACGA 780
Db 721 AGCAATATCTATAGACACACACAGCTCCCAATCATCAGGTGGGTAAACACTCCACACGA 780
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTGTGGTACCTGACCTATCATTTGTTG 840
Db 781 ATAGTCTGCTACCTCAGGAATATTTTGTGTGGTACCTGACCTATCATTTGTTG 840
QY 841 AATGCTCTTTCAGAAATCTATGCTTCTCTCATTTCTTAGTGCCCCCTATGACCATCTAC 900
Db 841 AATGCTCTTTCAGAAATCTATGCTTCTCTCATTTCTTAGTGCCCCCTATGACCATCTAC 900
QY 901 ACTGAACAAGATTTATACAAATCATGTCGTAACCTTAAGCCCCACAAACAAAGAGTACCCATT 960
Db 901 ACTGAACAAGATTTATACAAATCATGTCGTAACCTTAAGCCCCACAAACAAAGAGTACCCATT 960
QY 961 CTTCTTTTGTATCAGACAGAGGTGCTAGGACAGTAGGTACTGGCATTGGCAGTATC 1020
Db 961 CTTCTTTTGTATCAGACAGAGGTGCTAGGACAGTAGGTACTGGCATTGGCAGTATC 1020
QY 1021 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTCGACATGGAACAG 1080
Db 1021 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTCGACATGGAACAG 1080
QY 1081 GTCACTGACTCCCTGCTGCTTGCAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Db 1081 GTCACTGACTCCCTGCTGCTTGCAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT 1140
QY 1141 CAAATCGAAGAGCTTTAGACTGCTAACCGCCAAAGAGGGGAACTGTTTATTTT 1200
Db 1141 CAAATCGAAGAGCTTTAGACTGCTAACCGCCAAAGAGGGGAACTGTTTATTTT 1200
QY 1201 GGAAGAACGCTGTTTATTTATGTTAATCAATCAGAAATGTCATGAGAAAGTTAAAGAA 1260
Db 1201 GGAAGAACGCTGTTTATTTATGTTAATCAATCAGAAATGTCATGAGAAAGTTAAAGAA 1260
QY 1261 ATTGAGATCGAATACAATGTAGACAGAGAGCTTCAAAACACCCGACCTGGGGCTC 1320
Db 1261 ATTGAGATCGAATACAATGTAGACAGAGAGCTTCAAAACACCCGACCTGGGGCTC 1320
QY 1321 CTCAGCAATGGATGGCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 1321 CTCAGCAATGGATGGCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
QY 1381 TTACTCTCTTTGGACCTGTTTAACTCTCTGTTAAAGTTTGTCTCTTCCAGAAAT 1440
Db 1381 TTACTCTCTTTGGACCTGTTTAACTCTCTGTTAAAGTTTGTCTCTTCCAGAAAT 1440
QY 1441 GAAGCTGTAAGCTACAGATGCTTACAAATGGAACCCCA 1481
Db 1441 GAAGCTGTAAGCTACAGATGCTTACAAATGGAACCCCA 1481

```

; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US/09/890,340
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/IB00/00159
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: EP 99420041.8
; PRIOR FILING DATE: 1999-02-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 2030
; TYPE: DNA
; ORGANISM: MSRV-1 retrovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1626)
; US-10-637-565-18

Query Match      100.0%; Score 1481; DB 17; Length 2030;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60
DB 1 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60
QY 61 GCACCCCTCCATGCTGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAGA 120
DB 61 GCACCCCTCCATGCTGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAGA 120
QY 121 AGCGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCTCC 180
DB 121 AGCGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCTCC 180
QY 181 ACCTTCACTGCCACACACCATATGCCCCGCAACTGCTATAAATCTTGCACACTCTTGCATG 240
DB 181 ACCTTCACTGCCACACACCATATGCCCCGCAACTGCTATAAATCTTGCACACTCTTGCATG 240
QY 241 CATGCAATACTCATTTATGGACAGGGAATGATTAATCTTAGTTGCTCGAGGACTT 300
DB 241 CATGCAATACTCATTTATGGACAGGGAATGATTAATCTTAGTTGCTCGAGGACTT 300
QY 301 GGAGCCACTGCTGTTGGACTTACTTACCCCATACAGTATGCTGTAGTGGGGTGAAT 360
DB 301 GGAGCCACTGCTGTTGGACTTACTTACCCCATACAGTATGCTGTAGTGGGGTGAAT 360
QY 361 CAAGGTTCAGGCAAGAGAAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT 420
DB 361 CAAGGTTCAGGCAAGAGAAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT 420
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC 480
DB 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC 480
QY 481 CATACTCGCTGCTGAGCCTATTATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB 481 CATACTCGCTGCTGAGCCTATTATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
QY 541 CAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCACTTCCAGGCGCATACATTTCAATC 600
DB 541 CAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCACTTCCAGGCGCATACATTTCAATC 600
QY 601 CTGTTCTTGAACATGGAACACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTA 660
DB 601 CTGTTCTTGAACATGGAACACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTA 660
QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACCTCACCTGTGTAAATTT 720
DB 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACCTCACCTGTGTAAATTT 720
QY 721 AGCAATACTATAGACAAACAGGCTCCCAATGATGATGAGTGGGTAAACCTCCACACGA 780
DB 721 AGCAATACTATAGACAAACAGGCTCCCAATGATGATGAGTGGGTAAACCTCCACACGA 780

```

RESULT 4

```

US-10-416-642-4
; Sequence 4, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: ARVIZU, Chandra
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0842 PCT
; CURRENT APPLICATION NUMBER: US/10/416,642
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/249,407
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7477736CB1

```

```

781 ATAGTCTGCTACCTCAGGAATATTTTCTGTGTGTACTCTCAGCTATCATTTGTTG 840
DB 781 ATAGTCTGCTACCTCAGGAATATTTTCTGTGTGTACTCTCAGCTATCATTTGTTG 840
QY 841 AATGGCTCTTTCAGAACTATGCTGCTCTCTCATTTTAGTGCCCTTATGACCATCTAC 900
DB 841 AATGGCTCTTTCAGAACTATGCTGCTCTCTCATTTTAGTGCCCTTATGACCATCTAC 900
QY 901 ACTGAAACAAGATTTATACAATCATGCTAGTACCTTAAGCCCAACAACAAGAGTAGCCATT 960
DB 901 ACTGAAACAAGATTTATACAATCATGCTAGTACCTTAAGCCCAACAACAAGAGTAGCCATT 960
QY 961 CTTCTCTTTTATTCAGAGCAGGAGTGTAGGACAGACTAGTACTGGCAATTTGGCAGTATC 1020
DB 961 CTTCTCTTTTATTCAGAGCAGGAGTGTAGGACAGACTAGTACTGGCAATTTGGCAGTATC 1020
QY 1021 ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
DB 1021 ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
QY 1081 GTCACTGACTCCCTGGTCACTTGGCAAGTCAACTTAATCTCCCTAGCAGCAGTAGTCTCT 1140
DB 1081 GTCACTGACTCCCTGGTCACTTGGCAAGTCAACTTAATCTCCCTAGCAGCAGTAGTCTCT 1140
QY 1141 CAAATCGAAGAGCTTTTAGACTTGTCTAACCGCAAGAGGGGAACTGTTTATTTT 1200
DB 1141 CAAATCGAAGAGCTTTTAGACTTGTCTAACCGCAAGAGGGGAACTGTTTATTTT 1200
QY 1201 GGAGAGAACTGTTTATTTATTTAATCAATCCAGAAATTTGTCATGAGAAAGTTAAAGAA 1260
DB 1201 GGAGAGAACTGTTTATTTAATCAATCCAGAAATTTGTCATGAGAAAGTTAAAGAA 1260
QY 1261 ATTGAGATCGAATACATGTTAGCAGAGAGCTTCAAAACACGACGCTGGGGCTC 1320
DB 1261 ATTGAGATCGAATACATGTTAGCAGAGAGCTTCAAAACACGACGCTGGGGCTC 1320
QY 1321 CTCAGCCAAATGGATGCTGCTGCTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 1321 CTCAGCCAAATGGATGCTGCTGCTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATTG 1380
QY 1381 TTACTCTCTTTGGACCTGATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAAT 1440
DB 1381 TTACTCTCTTTGGACCTGATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAAT 1440
QY 1441 GAAGCTGTAAAGCTACAGATGCTCTTCAAAATGGAAACCCCA 1481
DB 1441 GAAGCTGTAAAGCTACAGATGCTCTTCAAAATGGAAACCCCA 1481

```

```

;
; NAME/KEY: unsure
; LOCATION: 1995
; OTHER INFORMATION: a, t, c, g, or other
; US-10-416-642-4

Query Match      93.8%; Score 1389.8; DB 17; Length 2074;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1424; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 1 ATGCCCTCCCTTATCATCTTTCTTTACTTGTCTTCTTACCCCTTTCGCTCTCACT 60
Db 64 ATGCCCTCCCTTATTTCTTTCTTTACTTGTCTTCTTACCCCTTTCGCTCTCACT 123
Qy 61 GCACCCCTCCATGCTCTGTAGAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
Db 124 GCACCCCTCCATGCTCTGTAGAACAGTAGCTCCCTTTACCAAGAGTTTCTATGGAGA 183
Qy 121 ACGGCGCTTCTTGAAATATTTGATGCCCCCATCATATATAGGAGTTTATCTAAGGGAATCTC 180
Db 184 ATGGCGGTCCGACATATTTGATGCCCCCATCGTATAGGAGTTTATCTAAGGGAATCTC 243
Qy 181 ACTTCTACTGCCACACCCATATATGCCCGGCAACTGTCTATTAATCTGCGCACTTTTGCAATG 240
Db 244 GCCTTCAACGCCACACCCATATATGCCCGGCAACTGTCTATTAATCTGCGCACTTTTGCAATG 303
Qy 241 CATGCAATACTCATTTATGACAGGGAATGATTAATCTAGTTGTCTGAGGAGCTT 300
Db 304 CATGCAATACTCATTTATGACAGGGAATGATTAATCTAGTTGTCTGAGGAGCTT 363
Qy 301 GGAGCCACTGTCTGTGAGCTTACTTCCACCCATACAGTAGTATGTCTGATGGGGTGGAAATT 360
Db 364 GGAGCCACTGTCTGTGAGCTTACTTCCACCCATACAGTAGTATGTCTGATGGGGTGGAGTT 423
Qy 361 CAAGGTGAGGCAAGAGAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGGAGAT 420
Db 424 CAAGATCAGGCAAGAGAAAACAAGTAAAGGAAGTAACTCTCCAACTGACCCGGGGTACAT 483
Qy 421 AGCACCCCTAGCCCTCAAGAGGACTAGTTCTCTCAAACTACATGAAGAACCTCCGTACC 480
Db 484 AGCACCCCTAGCCCTCAAGAGGACTAGTTCTCTCAAACTACATGAAGAACCTCCGTACC 543
Qy 481 CATACTCGCTGTGTGAGCTTATTAATACCAACCTCTACTCGGCTCCATGAGGTCTCAGCC 540
Db 544 CATACTCACCTGTGTGAGCTTATTAATACCAACCTCTACTCGGCTCCATGAGGTCTCAGCC 603
Qy 541 CAAACCCCTACTAACTGTGTGAGTGTGCTCCCTGCACTTCAGGCAATACATTTCAATC 600
Db 604 CAAACCCCTACTAACTGTGTGAGTGTGCTCCCTGCACTTCAGGCAATACATTTCAATC 663
Qy 601 CCTGTTCTGCAATGGAACAATTCAGCAGAGAAATTAACACCACTTCGGTTTGTAGTA 660
Db 664 CCTGTACTGAAACAATGGAATTAATTCAGCAGAGAAATTAACACCACTTCATTTTGTAGTA 723
Qy 661 GGACCTCTTGTTCCTCAATCTGAAATAACCCATACCTCAAACTCAGCTGTGTAATAATT 720
Db 724 GGACCTCTTGTTCCTCAATCTGGAATAATTCATACCTCAAACTCAGCTGTGTAATAATT 783
Qy 721 AGCAATACTATAGACACACAGCTCCCAATGCAATGAGGTGGGTAACACCTCCACACGA 780
Db 784 AGCAATACTATAGACACACACAGCTCCCAATGCAATGAGGTGGGTAACCTCCACACGA 843
Qy 781 ATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCTATCATTTGTTG 840
Db 844 ATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCTATCATTTGTTG 903
Qy 841 AATGGCTCTTCAGAACTATGTGCTCTCTCTCTTCTAGTGGCCCTTATGACCACTTAC 900
Db 904 AATGGCTCTTCAGAACTATGTGCTCTCTCTCTCTTCTAGTGGCCCTTATGACCACTTAC 963
Qy 901 ACTGAACAGAGTTTATCAATCATGTGTACCTAAGGCCCAACAACAAAGAGTAGTACCATT 960
Db 964 ACTGAACAGAGTTTATCAATCATGTGTACCTAAGGCCCAACAACAAAGAGTAGTACCATT 1023

;
; TITLE OF INVENTION: Methods and compositions for diagnosing
; and treating preeclampsia and gestational trophoblast
; FILE REFERENCE: GIN-600684
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930) ... (2546)
; US-09-902-535-1

Query Match      90.2%; Score 1335.4; DB 9; Length 2930;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATCTTTCTTTACTTGTCTTCTTACCCCTTTCGCTCTCACT 60
Db 930 ATGGCCCTCCCTTATCATCTTTCTTTACTTGTCTTCTTACCCCTTTCGCTCTCACT 989

```

QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAGA 120
DB 990 GCACCCCTCCATGCTGTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGAGA 1049
QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC 180
DB 1050 ATGAGCGTCCGGAATATTGATGCCCATCATATAGAGTTTCTAAGGGAACCCCC 1109
QY 181 ACCTTCACTGCCACACCCATATGCTGCCGCACTGCTATACTTCGCCACTCTTTGCATG 240
DB 1110 ACCTTCACTGCCACACCCATATGCTGCCGCACTGCTATACTTCGCCACTCTTTGCATG 1169
QY 241 CATGCAAACTCATATTATGGACAGGAAATGATTAATCTAGTTTCTCTGGAGGACTT 300
DB 1170 CATGCAAACTCATATTATGGACAGGAAATGATTAATCTAGTTTCTCTGGAGGACTT 1229
QY 301 GGAGCCACTGCTGTGGACTTACTTACCCTATACAGTATGCTGTGATGGGGGTGGAATT 360
DB 1230 GGAGTCACTGCTGTGGACTTACTTACCCTATACAGTATGCTGTGATGGGGGTGGAATT 1289
QY 361 CAAGGTCAAGCAAGGAAACAAAGTAAAGAGCAATCTCCAACTGACCCGGGGAAT 420
DB 1290 CAAGATCAGCAAGGAAACAAAGTAAAGAGCAATCTCCAACTGACCCGGGTACAT 1349
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
DB 1350 GGCACCTCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 1409
QY 481 CATACTCGCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
DB 1410 CATACTCGCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 1469
QY 541 CAAACCCCTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCAATACATTTCAATC 600
DB 1470 CAAACCCCTACTAACTGTTGGATGTCCTCCCTGAACTTCAGGCAATATGTTTCAATC 1529
QY 601 CTGTCTCTGAAACAATGGACACTTCAGACAGAAATAAACACCACTTCGTTTATGTA 660
DB 1530 CTGTCTCTGAAACAATGGACACTTCAGACAGAAATAAACACCACTTCGTTTATGTA 1589
QY 661 GGACCTCTGTTTCCAACTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 720
DB 1590 GGACCTCTGTTTCCAACTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 1649
QY 721 AGCAATATATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
DB 1650 AGCAATATATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 1709
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTTGTG 840
DB 1710 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTTGTG 1769
QY 841 AATGGCTCTTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCTATGACCATCTAC 900
DB 1770 AATGGCTCTTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCTATGACCATCTAC 1829
QY 901 ACTGAAACAAGATTTATCAATCATGTGCTACCTAAAGCCCAACAACAAGAGTACCATT 960
DB 1830 ACTGAAACAAGATTTATCAATCATGTGCTACCTAAAGCCCAACAACAAGAGTACCATT 1889
QY 961 CTTCCTTTTGTATCAGACAGGAGTCTTAGGAGACTAGGTACTGGCATTTGGCAGTATC 1020
DB 1890 CTTCCTTTTGTATCAGACAGGAGTCTTAGGAGACTAGGTACTGGCATTTGGCAGTATC 1949
QY 1021 ACAACCTCTACTAGTTTCTACTCAAACTATCTCAAGAAATAAATGTTGATGGAACAG 1080
DB 1950 ACAACCTCTACTAGTTTCTACTCAAACTATCTCAAGAAATAAATGTTGGAACAG 2009
QY 1081 GTCACTGACCTCCCTGCTCACTTGCAGATCACTTAACCTCCCTAGCAGAGTAGTCTT 1140
DB 2010 GTCCCGCACTCCCTGGTCACTTGTCAAGATCACTTAACCTCCCTAGCAGAGTAGTCTT 2069
QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTATTTT 1200

DB 2070 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACCTGTTATTTT 2129
QY 1201 GGAGAGAACGCTGTTATTTATTAATCAATCCAGAAATGTCTACTGAGAAAGTTAAAGAA 1260
DB 2130 GGGAAGAAATGCTGTTATTTATTAATCAATCCGGAATCGTCTACTGAGAAAGTTAAAGAA 2189
QY 1261 ATTGAGATCGAATACAAATGTAGAGAGAGAGCTTCAAAACACCGAAACGCTGGGCGCTC 1320
DB 2190 ATTGAGATCGAATACAAATGTAGAGAGAGAGCTTCAAAACACCGAAACGCTGGGCGCTC 2249
QY 1321 CTGAGCAATGAGTGGCTGCTTCCCTCTTCTAGGAGCTCTAGGAGCTCTAATATTG 1380
DB 2250 CTGAGCAATGAGTGGCTGCTTCCCTCTTCTAGGAGCTCTAGGAGCTCTAATATTG 2309
QY 1381 TTACTCTCTTTGAGACCTGATCTTTAACTCTCTTGTAAAGTTGTCTTCCAGAAAT 1440
DB 2310 CTACTCTCTTTGAGACCTGATCTTTAACTCTCTTGTAAAGTTGTCTTCCAGAAATC 2369
QY 1441 GAAGCTCTAAAGCTACAGATGGTCTTTACAAATGGAACCCCA 1481
DB 2370 GAAGCTCTAAAGCTACAGATGGTCTTTACAAATGGAACCCCAAGATGAGTCCAA 2410

RESULT 6
US-10-114-893-134
; Sequence 134, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: McVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-114-893-134

Query Match 90.2%; Score 1335.4; DB 13; Length 2946;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATCTTTCTTACTGTTCTTACCCCTTTTGGCTCTCACT 60
DB 928 ATGGCCCTCCCTTATCATATTTTCTTTTACTGTTCTTTTACCTCTTCACTCTCACT 987
QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAGA 120
DB 988 GCACCCCTCCATGCTGTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGAGA 1047
QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC 180
DB 1048 ATGAGCGTCCGGAATATTGATGCCCATCATATAGAGTTTCTTAAAGGGAACCCCC 1107
QY 181 ACCTTCACTGCCACACCCATATGCTGCCGCACTGCTATACTTCGCCACTCTTTGCATG 240


```
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCGTACC 480
Db 1348 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACC 1407
QY 481 CATACTCGCTGGTGAAGCTATTTAATACCAACCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 1408 CATACTCGCTGGTGAAGCTATTTAATACCAACCTCACTCGGCTCCATGAGGTCTCAGCC 1467
QY 541 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
Db 1468 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCACTTCAGGCCATATGTTTCAATC 1527
QY 601 CTGTTCTCTGAACAATGGAACAACTTCAGCAAGAAATAAACACACATTCGGTTTAGTA 660
Db 1528 CCGTACTCTGAACAATGGAACAACTTCAGCAAGAAATAAACACACATTCGGTTTAGTA 1587
QY 661 GGACCTCTGTTCCAACTCGGAATTAACCCATACCTCAAACTCAGCTGTGTAAATTT 720
Db 1588 GGACCTCTGTTTCCAACTCGGAATTAACCCATACCTCAAACTCAGCTGTGTAAATTT 1647
QY 721 AGCAATACTATAGACAAACACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCCAACGA 780
Db 1648 AGCAATACTATAGACAAACACAGCTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAA 1707
QY 781 ATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
Db 1708 ATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 1767
QY 841 AATGGCTCTCAGAACTATGCTTCCTCTCAATCTTAGTGCCCTATGACCATCTAC 900
Db 1768 AATGGCTCTCAGAACTATGCTTCCTCTCAATCTTAGTGCCCTATGACCATCTAC 1827
QY 901 ACTGAAACAAGATTTATACAACTCATGTCGTAACCTTAAGCCCAACAACAAGAGTACCCATT 960
Db 1828 ACTGAAACAAGATTTATACAACTATGTCATATCTAAGCCCGCAACAACAAGAGTACCCATT 1887
QY 961 CTTCTCTTTGTTATCAGACGAGGTGCTAGGACACTAGGTACTGGCAATGGCAGTATC 1020
Db 1888 CTTCTCTTTGTTATAGSAGCAGAGTGCTAGGTGCACTAGGTACTGGCAATGGCAGTATC 1947
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 1948 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 2007
QY 1081 GTCACCTGACTCCTCGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Db 2008 GTCGCGACTCCTCGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 2067
QY 1141 CAAAATCGAAGACTTTAGACTTGTAAACCGCCAAAGAGGGGGAACCTGTTTATTTTA 1200
Db 2068 CAAAATCGAAGACTTTAGACTTGTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTA 2127
QY 1201 GGAGAAGACGCTGTTATTTATGTTAATCAATCAGAAATGTCATCAGAGAAAGTTAAAGAA 1260
Db 2128 GGGGAAGATGCTGTTATTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGAA 2187
QY 1261 ATTGAGATCGAATACAAATGAGCAGAGAGGCTTCAAAAACACCGAAGCTGGGCGCTC 1320
Db 2188 ATTGAGATCGAATACAAATGAGCAGAGAGGCTTCAAAAACACCTGGACCTGGGCGCTC 2247
QY 1321 CTGAGCAATGGATGCCCTGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 2248 CTGAGCAATGGATGGCCCTGGATTTCTCCCTCTTAGGACCTCTAGCAGCTAATATTG 2307
QY 1381 TTACTCTCTTTGGACCTCTATCTTTAACTCCTCTGTTAAAGTTGTCTCTTCCAGAAAT 1440
Db 2308 CTACTCTCTTTGGACCTCTATCTTTAACTCCTCTGTTAACTTTGTCTCTTCCAGATC 2367
QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 2368 GAAGCTGTAAGCTACAAATGGAGCCCAAGATGAGTCCAA 2408
```

```
RESULT 8
US-10-133-036-3
; Sequence 3, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Human endogenous retrovirus
US-10-133-036-3

Query Match 90.1%; Score 1333.8; DB 17; Length 1617;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATATCTTTTCTTTACTGTTCTTTACCCCTTTTCGCTCTCACT 60
Db 1 ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTTACCCTCTTTTCACTCTCACT 60
QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 61 GCACCCCTCCATGCGCGTGTATGACCAGTAGTCCCTTACCAAGAGTTTCTATGGAGA 120
QY 121 ACGGCGCTTCCTGGAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGAACTCC 180
Db 121 ATGCAGCGTCCCGGAAATATTTGATGCCCATCTGATAGGAGTCTTTCTAAGGAACTCC 180
QY 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAAATCTGCCACTCTTTGCAATG 240
Db 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAAATCTGCCACTCTTTGCAATG 240
QY 241 CATGCAAACTACTTATTTGACAGGGGAAATGATTAATCTAGTTGCTCTGGAGACTT 300
Db 241 CATGCAAACTACTTATTTGACAGGGGAAATGATTAATCTAGTTGCTCTGGAGACTT 300
QY 301 GGAGCACTGTCTGTGGACTTACTTACCCTACCATGCTGTATGATGGGGGTGGAATT 360
Db 301 GGAGTCACTGTCTGTGGACTTACTTACCCTACCATGCTGTATGATGGGGGTGGAATT 360
QY 361 CAAGTCAAGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
Db 361 CAAGTCAAGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGTACAT 420
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCGTACC 480
Db 421 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGGTACC 480
QY 481 CATACTCGCTGGTGAAGCTATTTAATACCAACCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 481 CATACTCGCTGGTGAAGCTATTTAATACCAACCTCACTCGGCTCCATGAGGTCTCGGCC 540
QY 541 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
Db 541 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCACTTCAGGCCATATGTTTCAATC 600
QY 601 CCGTGTCTGAACAATGGAACAATCTCAGCAGAGAAATAAACACACATTCGGTTTAGTA 660
Db 601 CCGTGTCTGAACAATGGAACAATCTCAGCAGAGAAATAAACACACATTCGGTTTAGTA 660
QY 661 GGACCTCTGTTTCCAACTCGGAATTAACCCATACCTCAAACTCAGCTGTGTAAATTT 720
Db 661 GGACCTCTGTTTCCAACTCGGAATTAACCCATACCTCAAACTCAGCTGTGTAAATTT 720
```

```

Qy 721 AGCAATACATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
Db 721 AGCAATACATAGACACACACAGCTCCCAATGATCAGGTGGGTAACTCTCCACACAA 780
Qy 781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGGCTATCATTTGTTG 840
Db 781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGGCTATCATTTGTTG 840
Qy 841 AATGGCTCTTCCAGATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
Db 841 AATGGCTCTTCCAGATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
Qy 901 ACTGAAACAAGATTTATCAATCATGTGCTTAAAGCCCAACACAAAGAGTACCCATT 960
Db 901 ACTGAAACAAGATTTATACATGATGTATCATATCAAGCCCAACAAAGAGTACCCATT 960
Qy 961 CTTCCTTTTGTATCAGAGCAGGAGTCTAGGACAGCTAGTACTGAGCATTTGGCAGTATC 1020
Db 961 CTTCCTTTTGTATCAGAGCAGGAGTCTAGGACAGCTAGTACTGAGCATTTGGCAGTATC 1020
Qy 1021 ACAACCTCTACTAGTCTTACTACAACTATCTCAAGAAATAATGATGACATGGAACAG 1080
Db 1021 ACAACCTCTACTAGTCTTACTACAACTATCTCAAGAAATAATGATGGAACAG 1080
Qy 1081 GTCACTGACTCCCTGGTCACTTCAAGATCAATTAACCTCCCTAGCAGCAGTAGTCTT 1140
Db 1081 GTCCGCACTCCCTGGTCACTTCAAGATCAATTAACCTCCCTAGCAGCAGTAGTCTT 1140
Qy 1141 CAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACTGTATTTTAA 1200
Db 1141 CAAATCGAAGAGCTTTAGACTTCTAAACCGCTGAAAGAGGGGAACTGTATTTTAA 1200
Qy 1201 GGAGAAGACGCTGTTATTTATGTTAAATCAATCCAGATTTGCTACTGAGAAAGTTAAAGAA 1260
Db 1201 GGGGAAGAACTGCTGTTATTTATGTTAAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1260
Qy 1261 ATTCGAGATCGAATACATGATGAGCAGAGGAGCTTCAAAACACCGAAGCCTGGGGCTC 1320
Db 1261 ATTCGAGATCGAATACATGATGAGCAGAGGAGCTTCAAAACACCGAAGCCTGGGGCTC 1320
Qy 1321 CTCAGCCAAATGGATGCTTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 1321 CTCAGCCAAATGGATGCTTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Qy 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAAAGTTGTTCTCTCCAGAAATT 1440
Db 1381 CTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAAAGTTGTTCTCTCTCCAGAAATC 1440
Qy 1441 GAAGCTGTAAAGCTACAGATGCTTACAAATGGAAACCCCA 1481
Db 1441 GAAGCTGTAAAGCTACAGATGCTTACAAATGGAAACCCCA 1481

```

RESULT 9

US-09-854-867-21

; Sequence 21, Application US/09854867

; Publication No. US2003024356A1

; GENERAL INFORMATION:

; APPLICANT: JOAN, KNOLL H

; APPLICANT: ROGAN, PETER K

; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING

; FILE REFERENCE: 30307

; CURRENT APPLICATION NUMBER: US/09/854,867

; CURRENT FILING DATE: 2003-05-08

; NUMBER OF SEQ ID NOS: 613

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 21

; LENGTH: 8523

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: repeat_region

; LOCATION: (1)..(8523)

; OTHER INFORMATION: herv17

US-09-854-867-21

Query Match 90.1%; Score 1333.8; DB 10; Length 8523;

Best Local Similarity 93.8%; Pred. No. 0;

Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATATCTTTCTTTACTTGTCTCTTACCCCTTTCCGCTCTCACT 60

Db 5849 ATGGCCCTCCCTTATCATATCTTTCTTTACTTGTCTCTTACCCCTTTTACCCCTCTTCACTCTCACT 6908

Qy 61 GCACCCCTCCATGCTGTGTACAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120

Db 6909 GCACCCCTCCATGCTGTGTATGACAGTAGCTCCCTTTACCAAGAGTTTCTATGGAGA 6968

Qy 121 ACGGGCTTCTCGGAATATTGATGCCCATATAGGAGTTTATCTAAGGGAACCTCC 180

Db 6969 ATGACGGTCCCGGAATATTGATGCCCATATAGGAGTTTCTTCTAAGGGAACCTCC 7028

Qy 181 ACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATAAATCTGCACTCTTTGCAATG 240

Db 7029 ACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATAAATCTGCACTCTTTGCAATG 7088

Qy 241 CATGCAAAATCTCATTTATTGGACAGGGAATAATGATTAACTCTAGTTGTCTGAGGAGCTT 300

Db 7089 CATGCAAAATCTCATTTATTGGACAGGGAATAATGATTAACTCTAGTTGTCTGAGGAGCTT 7148

Qy 301 GGAGCACTGTCTGTTGGACTTACTTTCACCCATACAGATATGCTGATGGGGTGGAGTT 360

Db 7149 GGAGTCACTGTCTGTTGGACTTACTTTCACCCAACTGGTATGCTGATGGGGTGGAGTT 7208

Qy 361 CAAGGTCAGCAAGAGAAACAAAGTAAAGAAAGCAATCTCCCAACTGACCCCGGGGACAT 420

Db 7209 CAAGATCAGCAAGAGAAACAAAGTAAAGAAAGTAACTCTCCCAACTGACCCCGGGTACAT 7268

Qy 421 AGCACCCCTAGCCCCCTACAAAGGAGTGTCTCTCAAAACTACATGAAACCTCCCGTACC 480

Db 7269 GGCACCTCTAGCCCCCTAGCAAGGAGTGTCTCTCAAAACTACATGAAACCTCCCGTACC 7328

Qy 481 CATACTGCTGTGTAGCCTTATTAATACACCTCTACCTGGCTCCATGAGGTCTCAGCC 540

Db 7329 CATACTGCTGTGTAGCCTTATTAATACACCTCTACCTGGCTCCATGAGGTCTCAGCC 7388

Qy 541 CAATAACCTTAACTGTTGGATGTCTCCCTCTGCACTTCAGGCGCATACATTTCAATC 600

Db 7389 CAATAACCTTAACTGTTGGATGTCTCCCTCTGCACTTCAGGCGCATATGTTCAATC 7448

Qy 601 CCTGTTCTGAAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGTTTTAGTA 660

Db 7449 CCTGTACTGAAACAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGTTTTAGTA 7508

Qy 661 GGACCTCTTGTGTTCCAACTCTGGAATAACCCATACCTCAAACTCAGCTGTGTAATAATTT 720

Db 7509 GGACCTCTTGTGTTCCAACTCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATAATTT 7568

Qy 721 AGCAATACATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780

Db 7569 AGCAATACATAGACACACACAGCTCCCAATGATCAGGTGGGTAACTCTCCACACAA 7628

Qy 781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCTATCATTTGTTG 840

Db 7629 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCTATCATTTGTTG 7688

Qy 841 AATGGCTCTTCCAGATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900

Db 7689 AATGGCTCTTCCAGATCTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 7748

Qy 901 ACTGAACAAGATTTATCAATCATGCTACCTTAAGCCCAACCAAGAGTACCCATT 960

Db 7749 ACTGAACAAGATTTATCAATCATGCTACCTTAAGCCCAACCAAGAGTACCCATT 7808

Qy 961 CTTCCTTTTGTATCAGAGCAGGAGTGTCTAGGACAGCTAGGTACTGCGATTGGCAGTATC 1020

Db 7809 CTTCTTTTGTATAGGACGAGTGTAGTGCACCTAGTACTGGCATGGCGGTATC 7868
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATATGGTGACATGGAACAG 1080
Db 7869 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATAATGGGACATGGAACGG 7928
Qy 1081 GTCACTGACTCCTGGTCACTTGCAGATCACTTAACCTCCTAGCAGCAGTAGTCCTT 1140
Db 7929 GTGCGCGACTCCTGGTCACTTGCAGATCACTTAACCTCCTAGCAGCAGTAGTCCTT 7988
Qy 1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGGAACCTGTTTATTTTA 1200
Db 7989 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGGAACCTGTTTATTTTA 8048
Qy 1201 GGAGAGAAACGCTGTTATTTATGTTAAATCAATCCAGAAATGTCACGTAGAGAAAGTTAAAGAA 1260
Db 8049 GGGGAAGAAATGCTGTTATTTATGTTAAATCAATCCGGAATCGTCACGTAGAGAAAGTTAAAGAA 8108
Qy 1261 ATTGAGATCGAATCAATGTAGACAGAGAGCTTCAAAACACCGAGCTGGGGCCTC 1320
Db 8109 ATTGAGATCGAATCAACGTTAGACAGAGAGCTTCAAAACACCTGGACCTGGGGCCTC 8168
Qy 1321 CTCAGCAATGGATGGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTTG 1380
Db 8169 CTCAGCAATGGATGGCCCTGGATTTCTCCCTCTTTAGGACCTCTAGCAGCTAATATTTG 8228
Qy 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTTAAAGTTGTCCTTCGAGAAAT 1440
Db 8229 CTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTTAACTTTGTCTCTTCCAGAAATC 8288
Qy 1441 GAAGCTGAAGCTACAGATGCTTTACAAATGGAAACCCCA 1481
Db 8289 GAAGCTGTAATACTCAAAATGGAGCCCAAGATGCAGTCCAA 8329

RESULT 10

US-09-873-367C-81
; Sequence 81, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-81
Query Match 90.1%; Score 1333.8; DB 10; Length 56093;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGCGCCCTCCCTTATCATACTTTTCTTCTTACTGTTCTTCTTACCCCTTTGGCTCTCACT 60

Db 35879 ATGCGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTTACCCTCTTTTCACTCTCACT 35938
Qy 61 GCACCCCTCATGCTGCTGTACAAACAGTAGTCCCTTACCAGAGAGTTTCTTATGAAGA 120
Db 35939 GCACCCCTCATGCTGCTGTATGACAGAGTAGTCCCTTACCAGAGAGTTTCTTATGAGA 35998
Qy 121 ACGGGGCTTCTGCGAAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 35999 ATGACAGCTCCCGAATATTTGATGCCCATCATATAGGAGTTTCTTCTAAGGGAACCCCC 36058
Qy 181 ACCTTCATGCGCACACCCCATATGCGCCCGCACTGCTATTAACCTCTGCCACTCTTTGCAATG 240
Db 36059 ACCTTCATGCGCACACCCCATATGCGCCCGCACTGCTATTAACCTCTGCCACTCTTTGCAATG 36118
Qy 241 CATCAATATCTCAATTTTGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT 300
Db 36119 CATCAATATCTCAATTTTGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT 36178
Qy 301 GGAGCCACTGCTCTGTTGGACTTACTTCAACCATACCACTATGCTGTATGCGGGTGGAAAT 360
Db 36179 GGAGCCACTGCTCTGTTGGACTTACTTCAACCATACCACTATGCTGTATGCGGGTGGAGTT 36238
Qy 361 CAAGGTGAGCAAGAGAAACCAAGTAAGAGAACTCTCCCACTGACCCGGGACAT 420
Db 36239 CAAGGTGAGCAAGAGAAACCAAGTAAGAGAACTCTCCCACTGACCCGGGACAT 36298
Qy 421 AGCACCCCTAGCCCTTCAAAAGGACTAGTTCTCTCAAACTACATGAAACCTCCGTACC 480
Db 36299 GGCACCTCTAGCCCTTCAAAAGGACTAGTTCTCTCAAACTACATGAAACCTCCGTACC 36358
Qy 481 CATACTCGCTGCTGAGCCTATTTAATAACCACTCTCGGCTCCATGAGGTCTCAGCC 540
Db 36359 CATACTCGCTGCTGAGCCTATTTAATAACCACTCTCGGCTCCATGAGGTCTCAGCC 36418
Qy 541 CAAACCCCTACTAATGTTGGATGCTCTCCCTGCACTTTCAGGCCATACATTTCAATC 600
Db 36419 CAAACCCCTACTAATGTTGGATGCTCTCCCTGCACTTTCAGGCCATATGTTTCAATC 36478
Qy 601 CCTGTTCTGAACTAATGGAACAACTTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 660
Db 36479 CCTGTTCTGAACTAATGGAACAACTTTCAGCACAGAAATAAACACCACTTCCGTTTTAGTA 36538
Qy 661 GGACCTCTGTTTCCAATCTCGAAATAACCATACCTCAAACTCCTGCTGTGTAATTT 720
Db 36539 GGACCTCTGTTTCCAATCTCGAAATAACCATACCTCAAACTCCTGCTGTGTAATTT 36598
Qy 721 AGCAATCTATAGACACACCACTCCCAATGCACTCAGTGGGTAAACCTCCCAACAGA 780
Db 36599 AGCAATCTATAGACACACCACTCCCAATGCACTCAGTGGGTAAACCTCCCAACAGA 36658
Qy 781 ATAGTCTGCTACCCCTCAGGAATATTTTGTGTGTGTACCTCAGCCTATCATTTGTTG 840
Db 36659 ATAGTCTGCTACCCCTCAGGAATATTTTGTGTGTGTACCTCAGCCTATCATTTGTTG 36718
Qy 841 AATGGCTCTCAGAACTATGTTGCTCTCTCATTTCTAGTGCCCTCTATGACCATCTAC 900
Db 36719 AATGGCTCTCAGAACTATGTTGCTCTCTCATTTCTAGTGCCCTCTATGACCATCTAC 36778
Qy 901 ACTGAAACAGATTTTATACAATCATGTCCTAGCCCAACCAAAAGAGTACCCATT 960
Db 36779 ACTGAAACAGATTTTATACAATCATGTCCTAGCCCAACCAAAAGAGTACCCATT 36838
Qy 961 CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGCAGACTAGGTACTGGCATTTGGCAGTATC 1020
Db 36839 CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGTGCACTAGGTACTGGCATTTGGCAGTATC 36898
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATATGGTGACATGGAACAG 1080
Db 36899 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTATATATGGGACATGGAACAG 36958
Qy 1081 GTCACTGACTCCTCGTCACTCTGCAAGATCAACTTAATCTCCCTAGCAGCAGTAGTCTCT 1140

Db 36959 GTCCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAATCCCTAGCAGCAGTAGTCCTT 37018
Qy 1141 CAAATCGAAGACTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTTA 1200
Db 37019 CAAATCGAAGACTTTAGACTTGTCTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTA 37078
Qy 1201 GGAGAAGAGCGCTGTTTATTTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
Db 37079 GGGGAAGATGCTGTTTATTTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 37138
Qy 1261 ATTCGAGATCGAATCAATGATAGAGCAGAGGAGCTTCAAAACCGAAGCTGGGCGCTC 1320
Db 37139 ATTCGAGATCGAATCAACAGCTAGAGCAGAGGAGCTTCGAAACACTGGACCCCTGGGCGCTC 37198
Qy 1321 CTGAGCCATGGAATGCCCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAATATG 1380
Db 37199 CTCAGCAATGGAATGCCCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAATATG 37258
Qy 1381 TTACTCTCTTTGAGCCCTGATCTTTAACTCTCTTGTAAAGTTTGTCTTCCAGAAAT 1440
Db 37259 CTACTCTCTTTGAGCCCTGATCTTTAACTCTCTTGTAAAGTTTGTCTTCCAGAAAT 37318
Qy 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 37319 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAA 37359

RESULT 11
US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauclia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-30

Query Match 89.4%; Score 1324.2; DB 17; Length 7582;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 1365; Conservative 24; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATACTTTTCTTTACTGTTCTTACCCCTTTTCGCTCCTCACT 60
Db 5581 ATGGSCCTCCCTTATCATATTTTCTTASGTSTTTTACCTSTTTCACTCCTCACT 5640

Qy 61 GCACCCCTCCATGCTGCTGTACACACAGTAGTCTCCCTTACCAAGAGTTTCTTGAAGA 120
Db 5641 GCACCCCTCCATGCTGCTGTATGACACAGTAGTCTCCCTTACCMAGAGTTTCTTATGAGA 5700

Qy 121 ACGGGCTTCTCGGAATATTGATGCCCATATATAGAGTTTATCTAAGGGAAACTCC 180
Db 5701 ATGCAGCGTCCCGGAATATTGATGCCCATATATAGAGTTTCTTSTAAGGAAACCC 5760

Qy 181 ACCTTCACTGCCACACCCATATGCTCCGCAACTGCTATAAATCTGCACTCTTTGCAATG 240
Db 5761 ACCTTCACTGCCACACCCATATGCTCCGCAACTGCTATAAATCTGCACTCTTTGCAATG 5820

Qy 241 CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCGGAGGACTT 300
Db 5821 CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTCGGAGGACTT 5880

Qy 301 GGAGCCACTGCTGTTGGACTTACTTACCCATACAGTATGCTGATGGGGTGGAAAT 360
Db 5881 GGAGTCACTGCTGTTGGACTTACTTACCCAACTGATGATGCTGATGGGGTGGAGTT 5940

Qy 361 CAAGGTCAGGCAAGAGAAACAAAGTAAGAAAGAGCAATCTCCCAACTGACCCCGGGGACAT 420
Db 5941 CAAGATCAGGCAAGAGAAACAAAGTAAGAAAGTAATCTCCCAACTGACCCGGGTACAT 6000

Qy 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCGTACC 480
Db 6001 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACC 6060

Qy 481 CATACTCGCTGGTGGCTTATTTAATACCACTCTACCTGGCTCCATGAGGTCTCAGCC 540
Db 6061 CATACTCGCTGGTGGCTTATTTAATACCACTCTACCTGGCTCCATGAGGTCTCAGCC 6120

Qy 541 CAAAACCCCTACTAACTGTTGGATGCTCCCTCGCACTTTCAGGCGCATACATTTCAATC 600
Db 6121 CAAAACCCCTACTAACTGTTGGATGCTCCCTCGCACTTTCAGGCGCATATGTTTCAATC 6180

Qy 601 CCTGTTCTGAAACAATGGAACAACCTTCCAGCAGAAATAACACCACTTCGTTTATGTA 660
Db 6181 CCTGTACTGAAACAATGGAACAACCTTCCAGCAGAAATAACACCACTTCGTTTATGTA 6240

Qy 661 GGACCTCTTCTTCCATCTGGAATAACCACTCAACCTCAACCTCTGTTGTAATTT 720
Db 6241 GGACCTCTTCTTCCATCTGGAATAACCACTCAACCTCAACCTCTGTTGTAATTT 6300

QY 1021 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGGTGACATGGACAG 1080
 DB 1783 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGGTGACATGGACAG 1842
 QY 1081 GTCACTGACCTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1140
 DB 1843 GTGCCGACCTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1902
 QY 1141 CAANAATCGAAGAGCTTTAGACTTGTCAACCGCCAAAGAGGGGGAACCTGTTTATTTT 1200
 DB 1903 CGAATCGAAGAGCTTTAGACTTGTCAACCGCTGAGAGGGGGAACCTGTTTATTTT 1962
 QY 1201 GGAAGAAACGCTGTTTATTTATTTAATCAATCAGAAATTTCTACTGAGAAAGTTAAGAA 1260
 DB 1963 GGGGAAGATGCTGTTTATTTAATCAATCGGAATCTCTACTGAGAAAGTTAAGAA 2022
 QY 1261 ATTCCAGATCGAATACAAATAGAGCAGAGAGCTTCAAAACCCGAAACCTGGGGCTC 1320
 DB 2023 ATTCCAGATCGAATACAAATAGAGCAGAGAGCTTCAAAACCCGAAACCTGGGGCTC 2082
 QY 1321 CTGAGCAATGGATGCGCTGGTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATTG 1380
 DB 2083 CTCAGCCGATGGATGCGCTGGTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATTG 2142
 QY 1381 TTACTCTCTTTGGACCTGATCTTTAACTCTCTTAAAGTTGTTCTCTCCAGAAAT 1440
 DB 2143 CTACTCTCTTTGGACCTGATCTTTAACTCTCTTAAAGTTGTTCTCTCCAGAAAT 2202
 QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 DB 2203 GAAGCTGTAAGCTACAAATGGAGCCCAAGATGCAATGCCAA 2243
 RESULT 13
 US-10-632-793-26
 ; Sequence 26, Application US/10632793
 ; Publication No. US20040048298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PARANHOS-BACCALA, Glauca
 ; APPLICANT: MALLET, Francois
 ; APPLICANT: VOISSET, Cecile
 ; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
 ; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
 ; FILE REFERENCE: 110048
 ; CURRENT APPLICATION NUMBER: US/10/632,793
 ; CURRENT FILING DATE: 2003-08-04
 ; PRIOR APPLICATION NUMBER: US/09/869,927
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: PCT/FR00/00144
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR APPLICATION NUMBER: FR 99/00888
 ; PRIOR FILING DATE: 1999-01-21
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 26
 ; LENGTH: 2782
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-632-793-26
 Query Match 88.2%; Score 1306.6; DB 17; Length 2782;
 Best Local Similarity 92.6%; Pred. No. 0;
 Matches 1372; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATCATCTTTCTTTTACTGTTCTTTACCCCTTTTCGCTCTCACT 60
 DB 763 ATGGCCCTCCCTTATCATCTTTCTTTTCTGTTGTTCTTTACCCCTTTTCGCTCTCACT 822
 QY 61 GCACCCCTCCATGCTGTGTAACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
 DB 823 GCACCCCTCCATGCTGTGTAACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 882
 QY 121 ACGGGCTCTCTGGAATATTGATGCCCCCATATAGGAGTTTATCTAAGGGAACTCC 180

DB 883 ATGACGCGTCCCGGAATAATTGATGCCCATCGTAGGAGTCTTTCTAAGGGAACCCCC 942
 QY 181 ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATAAATCTCTCCACTCTTTTGCATG 240
 DB 943 ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATAAATCTCTCCACTCTTTTGCATG 1002
 QY 241 CATGCAAAATACCTATTATTGGACAGGAAATGATTAACTCTAGTTGCTGCTGAGACATT 300
 DB 1003 CATGCAAAATACCTATTATTGGACAGGAAATGATTAACTCTAGTTGCTGCTGAGACATT 1062
 QY 301 GAGGCCACTGCTGTTGGACTTACTTTCACCCATACAGTATGCTCTCATGGGGTGGAAATT 360
 DB 1063 GAGGTCACTGCTGTTGGACTTACTTTCACCCAACTGGTATGCTCATGGGGTGGAGATT 1122
 QY 361 CRAAGTCAGGCAAGAGCAAAAAAAGTAAAGAGCAATCTCCCAACTGACCCCGGGGACAT 420
 DB 1123 CRAAGTCAGGCAAGAGCAAAAAAAGTAAAGAGCAATCTCCCAACTGACCCCGGGGACAT 1182
 QY 421 AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCCTCCGTAAC 480
 DB 1183 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTAAC 1242
 QY 481 CATATCGCTGCTGAGCCCTATTATAACACCTCATCTCGGCTCCATGAGGTCTCAGCC 540
 DB 1243 CATATCGCTGCTGAGCCCTATTATAACACCTCATCTCGGCTCCATGAGGTCTCAGCC 1302
 QY 541 CAAAACCTCTAACTGTTGGATGCTCCCTCCCTGACCTTCAGGGCCATACATTTCAATC 600
 DB 1303 CAAAACCTCTAACTGTTGGATGCTCCCTCCCTGACCTTCAGGGCCATACATTTCAATC 1362
 QY 601 CTGTTCTCTGAAACAATGGAAACAATTTAGCACAGAAATAAACAACCACTTCCTGTTTATGTA 660
 DB 1363 CTGTTCTCTGAAACAATGGAAACAATTTAGCACAGAAATAAACAACCACTTCCTGTTTATGTA 1422
 QY 661 GGAACCTCTGTTTCCAACTCTGAAATAAACCATACTCAAACTCACTCTGTGTAAATTTT 720
 DB 1423 GGAACCTCTGTTTCCAACTCTGAAATAAACCATACTCAAACTCACTCTGTGTAAATTTT 1482
 QY 721 AGCAATACTATAGACACCAACCTCCCAATGATCAGTGGGTAAACACCTCCACACGA 780
 DB 1483 AGCAATACTATAGACACCAACCTCCCAATGATCAGTGGGTAAACACCTCCACACGA 1542
 QY 781 ATAGTCTGCTCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCTATCATTTGTTG 840
 DB 1543 ATAGTCTGCTCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCTATCATTTGTTG 1602
 QY 841 AATGGCTCTTCAGAAATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 DB 1603 AATGGCTCTTCAGAAATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1662
 QY 901 ACTGAAACAAGATTATACAACTCATGCTGTAACCTTAAGCCCAACCAAGAGTACCCATT 960
 DB 1663 ACTGAAACAAGATTATACAACTCATGCTGTAACCTTAAGCCCAACCAAGAGTACCCATT 1722
 QY 961 CTTCCTTTTCTATCAGAGCAGAGTCTAGGACAGCTAGGTACTGTCGATTTGGCAGTATC 1020
 DB 1723 CTTCCTTTTCTATCAGAGCAGAGTCTAGGACAGCTAGGTACTGTCGATTTGGCAGTATC 1782
 QY 1021 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAAATGGTGACATGGAAACAG 1080
 DB 1783 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAAATGGTGACATGGAAACAG 1842
 QY 1081 GTCACTGACCTCCCTGGTCACTTGTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1140
 DB 1843 GTGCCGACCTCCCTGGTCACTTGTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1902
 QY 1141 CAANAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTT 1200
 DB 1903 CGAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGGGGGAACCTGTTTATTTT 1962
 QY 1201 GGAAGAAACGCTGTTTATTTATTTAATCAATCAGAAATTTGCTACTGAGAAAGTTAAGAA 1260

Db 1963 GGGGAAGATGCTGTTATTATGTTAAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGA 2022

Qy 1261 ATTGAGATCGAATACAATGTAGACAGAGGAGCTTTCAAAACACCGAAGCGTGGGGCTTC 1320

Db 2023 ATTCCAGATCGAATACAACGTATAGCAGAGGAGCTTCGAAACACATGGACCCCTGGGGCTTC 2082

Qy 1321 CTCAGCAATGATGCCCTGGGTTCTCCCTTCTTTAGGACCTCTAGCAGCTCTAATATTG 1380

Db 2083 CTCAGCCGATGATGCCCTGGGATTCTCCCTTCTTTAGGACCTCTAGCAGCTAATAATTG 2142

Qy 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCCTTGTAAAGTGTGTCTTTCCAGAAAT 1440

Db 2143 CTACTCCTCTTTGGACCCCTGTATCTTTGACCTCCTTGTAACTTTGTCTTCCAGAAATC 2202

Qy 1441 GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAAACCCCA 1481

Db 2203 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAATGCCAA 2243

RESULT 14

US-10-087-192-730/c

; Sequence 730, Application US/10087192

; Publication No. US20020182586A1

; GENERAL INFORMATION:

; APPLICANT: Morris, David W.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

; TITLE OF INVENTION: CANCER

; FILE REFERENCE: 529452000122

; CURRENT APPLICATION NUMBER: US/10/087,192

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 2059

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 730

; LENGTH: 161334

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)...(161334)

; OTHER INFORMATION: n = A, T, C or G

US-10-087-192-730

Query Match 86.3%; Score 1277.4; DB 13; Length 161334;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1399; Conservative 0; Mismatches 81; Indels 33; Gaps 4;

Qy 1 ATGGCCCTCCCTATCATACTTTTCTCTTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60

Db 102037 ATGGCCCTCCCTATCATATTTTCTCTTTACTGTTCTCTTACCCCTTTTCACTCTCACT 101978

Qy 61 GCACCCCTCCATGCTGCTACAAACGAGTAGCTCCCTTACCAGAGTTCTCTATGAAGA 120

Db 101977 GCACCCCTCCATGCTGCTACAAACGAGTAGCTCCCTTACCAGAGTTCTCTATGAAGA 101918

Qy 121 ACAGCGCTTCTCGAAATATTTAGTGTGCCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180

Db 101917 ATGGCGCTTCCCAGAAATATTGATGCCCCATCAATAGGAGTTTACCTAAGGAACTCC 101858

Qy 181 ACCTTCAGTCCGCACACCCATATGCCCCGCACTGTCTAATCTCTGCCACTCTTTGCATG 240

Db 101857 ACCTTCAGTCCGCACACCCATATGCCCCCACTGTCTAATCTCTGCCACTCTTTGCATG 101798

Qy 241 CATGCAAACTCTCATTTATGGACAGGAAATGATTAATCTCTAGTTGTCTGAGGACTT 300

Db 101797 CATGCAAACTCTCATTTATGGACAGGAAATGATTAATCTCTAGTTGTCTGAGGACTT 101738

Qy 301 GGAGCCACTGTCTGTGGACTTACTTTACCCATACCAGATATGTTCTGATGGGGTGGAAAT 360

Db 101737 GGAGCCACTGTCTGTCGAGCTTACTTCCACCATACTGGTATGTCTGAGGGGGTGGAGTT 101678

Qy 361 CAAGGTGAGCAAGAGAAACAAAGTAAGAGCAATCTCCCACTGACCCGGGAGCAT 420

Db 101677 CAAGATCAGGCAAGAGAAACAAATGTAAAGAAATTAATCTCCCACTGACCCGGGTAGCAT 101618

Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTC-AAAACTACATGAAACCCCTCCGTAC 479

Db 101617 AGCACCCCTAGCCCTACAAAGGACTAGATCTCTCTAAAACTACATGAAACCCCTCCATAC 101558

Qy 480 CCATACCTGCGCTGTGAGCCCTATTAAATACACCCCTCACTCGGTCCATAGGCTCTCAGC 539

Db 101557 CCATACCTGCGCTGTGAGCCCTATTAAATACACCCCTCACTCGGTCCATAGGCTCTCAGC 101498

Qy 540 CCAAAACCCCTACTAATCTTGGATGTGCTCCCTGACCTTTCAGGCCATACATTTCAAT 599

Db 101497 CCAAAACCCCTACTAATCTTGGATGTGCTCCCTGACCTTTCAGGCCATGCAATTTCAAT 101438

Qy 600 CCCTGTTCTCGAAACAATGGAACAACCTTCAGCACAGAAATAAACACCACTTCCTGTTTTAGT 659

Db 101437 CCCTGTACTCGAACAATGGAACAACCTTCAGCACAGAAATAAACACCACTTCCTGTTTTAGT 101378

Qy 660 AGGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAAACTCCTGCTGTGTTAAAT 719

Db 101377 AGGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAAACTCCTGCTGTGTTAAAT 101318

Qy 720 TAGCAATACTATAGCACAAACAGCTCCCAATGCACTAGGTGGGTAAACACCTCCACACAG 779

Db 101317 TAGCAATACTATAGCACAAACAGCTCCCAATGCACTAGGTGGGTAACTCTCTCCACACAG 101258

Qy 780 AATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACCTCAGGCTATCATTTGTTT 839

Db 101257 AATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACCTTAGCCTATCGTTGTTT 101198

Qy 840 GAATGGCTCTTCAGAAATCTATGTGCTCTCTCTCATTTAGTGCCCCCTATGACCATCTA 899

Db 101197 GAATGGCTCTTCAGAAATCTATGTGCTCTCTCTCATTTAGTGCCCCCT-ATGACCATTTA 101139

Qy 900 CACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCAACAAACAAAGAGTACCCAT 959

Db 101138 CACTGAACAAGATTTATACAATCATGTCGTACCTAAGCCCAACAAACAAAGAGTACTCAT 101079

Qy 960 TCTTCTCTTTGTTATCAGACGAGGAGTCTAGGACAGCTAGGTACTGCGCTTGGCAGTAT 1019

Db 101078 TCTTCTCTTTGTTATCAGACGAGGAGTCTAGGTGAGCTAGGTCTGCGCTTGGCAGTAT 101019

Qy 1020 CACAACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTATCTCAAGAACTCAATGGATG 1079

Db 101018 CACAACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTCAATGGATG 100959

Qy 1080 GGTCACTGACTCCCTGGTCCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1139

Db 100958 GGTGCCGACTCCCTGGTCCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 100899

Qy 1140 TCATAATCGAAGAGCTTTAGACTTGCCTAACCGCCAAAGAGGGGAACTGTTTATTTT 1199

Db 100898 TCATAATCGAAGAGCTTTAGACTTGCCTAACCGCTTGCCTTGCCTTGCCTTGCCTTGCCT 100839

Qy 1200 AGGAGAAAGACGCTG-----TTATTATGTTAAATCA 1229

Db 100838 AGGGGAAGAATGTTGTTATTATTTAGCGGAAGAATGTTGTTATTATGTTAAATCA 100779

Qy 1230 ATCCAGAAATGTGCTAGAGAAAGTTAAAGAAATTCGAGATCGAATACGAATAGTAGAGCAGA 1289

Db 100778 ATCCGAAATGTGCTAGAGAAAGTTGAAGAAATTCGAGATCGAATAGTAGAGCAGA 100719

Qy 1290 GGAGCTTC-AAAAACCGAAGCCTGGGCTCCTCAGCCCAATGAGATGCCCTGGGTTCTCC 1348

Db 100718 GGAGCTTCAAAAAACACACAGACCTGGGCTCCTCAGCCCAATGAGATGCCCTGGGTTCTCC 100659

Qy 1349 CCTTCTTAGGACTCTAGCAGCTCTAATATTTGTTACTCTCTTTCGACCCCTGATCTTTA 1408

Db 100658 CCTTCTTAGGACTCTAGCAGCTCTAATATTTGTTACTCTCTTTCGACCCCTGATCTTTA 100599

QY 1409 ACCTCTTGTAAAGTTTCTCTTCCAGAAATGAAGCTGTAAGAGCTACAGATGGTCTTAC 1468
Db 100598 ACTCTCTTGTAAAGTTTCTCTTCCAGAAATGAAGCTGTAAGAGCTACAGATGGTCTTAC 100539
QY 1469 AAATGGAACCCCA 1481
Db 100538 AAATGGAACCCCA 100526

RESULT 15
US-09-864-761-4444
; Sequence 4444, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4444
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002345.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
US-09-864-761-4444
Query Match 76.4%; Score 1131.6; DB 9; Length 1894;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;
QY 164 TATCTAAGGGAACCTCCACCTTCACTGCCCACACCCATATGCCCAGCACTGCTATAACT 223
Db 1 TATCTAAGGGAACCTCCACCTTCACTGCCCACACCCATATGCCCAGCACTGCTATAACT 60
QY 224 CTGCCACTCTTTGTCATGTCATGCAATATCTCATTTATTTGGACAGGGAATAATGATTAACTCTA 283
Db 61 CTGCCACTCTTTGTCATGTCATGCAATATCTCATTTATTTGGACAGGGAATAATGATTAACTCTA 120
QY 284 GTTGTCTGTGAGGACTTGGAGCCACTGTCTGTGTGGACTTACTTCCACCATACCAATATGT 343
Db 121 GTTGTCTGTGAGGACTTGGAGCCACTGTCTGTGTGGACTTACTTCCACCATACCAATATGT 167
QY 344 CTGATGGGGTGGAAATTCAGGTTCAGCAAGAGAAAACAAGTAAAGAGAGCAATCTCCC 403
Db 168 CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAACTCTCCC 227
QY 404 AACTGACCCGGGACATAGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTAC 463
Db 228 AACTGACCTGGGTACATAGCACCCCTGGCCCTTACAAAGGACTAGTCTCTCAAAACTAC 287
QY 464 ATGAAACCTTCCGTACCCATCTCGCTGTGTAGCCCTTATTAATACACCTCTACTCGGC 523
Db 288 ATGAAACCTTCCATACCCATCTCGCTGTGTAGCCCTTATTAATACACCTCTACTCGGC 347
QY 524 TCCATGAGGTCTCAGCCCAAAACCTTACTTACTTGGATGTGCTGCTCCCTGACACTCA 583
Db 348 TCCATGAGGTCTCAGCCCAAAACCTTACTTACTTGGATGTGCTGCTCCCTGACACTTCA 407
QY 584 GGCCATACATTTCAATCCCTGTTCTGAAACAATGGAACAACCTTCAGCACAGAAATAACA 643
Db 408 GGCCATACATTTCAATCCCTTATCTGAAACAATGGAACAACCTTCAGCACAGAAATAACA 467
QY 644 CCACCTTCCGTTTGTAGTAGGACCTCTTGTGTTTCCAAATCTGGAATAACCCATACCTCA 703
Db 468 CCACCTTCTGTTTGTAGTAGGCTCTC---TTTCCAATCTGGAATAACCCATACCTCA 524
QY 704 TCACCTGTGTAATAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCAATCAGGTGG 763
Db 525 TCACCTGTGTAATAATTTAGCAATACTATAGACACAGCCAACTCCCAATGCAATCAGGTGG 584
QY 764 TAAACCTCCACACAGAAATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACTCT 823
Db 585 TAACTCTCCACACAGAAATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACTCT 644
QY 824 CAGCCTATCATTTGTTGAATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGC 883
Db 645 CAGCCTATCATTTGTTGAATGGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGC 704
QY 884 CCCCTATGACATCTACACTGAAACAAGATTTATCAATCATCTGCTACCTTAAGCCCCACA 943
Db 705 CCCCTATGACATCTACACTGAAACAAGATTTATCAATCATCTGCTACCTTAAGCCCCACA 764
QY 944 ACAAAAGATGACCAATCTTCTTCTTCTTATCAGACAGGAGTCTAGGAGAGCTAGGTA 1003
Db 765 ACAAAAGATGACCAATCTTCTTCTTCTTATTTGGAGCAGGAGTCTAGGCGGAGTAGCTA 824
QY 1004 CTGGCATTGGCAGTATCACAACTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAA 1063
Db 825 CTGGCATTGGGCGTATCACAACTCTACTCAGTTTCTACTACAACTGTCTCAAGAAATAA 884
QY 1064 ATGGTGATGGAACAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123

Db	885	ATGGTGACATGGAATGGGTGCGCTGATACCCCTGTGCACCTTTGCGAAGATCAACTTAACCTCCC	944
Qy	1124	TAGCAGCAGTAGTCCTTTCAAAATCGAAGAGCTTTAGACTTGCTTAACCGCCAAAGAGGGG	1183
Db	945	TAGCAGCAGTAGTCCTTTCAAAATCGAAGAGCTTTAGACTTGCTTAACCGCGGAAGCGGG	1004
Qy	1184	GAACCTGTTTATTTTATAGGAGAGAAGCGCTGTTATTATGTTAATCAATCGAGAAATTTGTCA	1243
Db	1005	GAACCTTTTATTTTATAGAGAAATATGCTGTGTGTTATGTTAATCAATCGGAATCATCA	1064
Qy	1244	CTGAGAAAGTTAAAGAAATTCGAGATCGAATCAAACTAGACAGAGAGAGCTTCAAAACA	1303
Db	1065	CCGAGAAAGTTAAAGAAATTCAGGTCGAATATACGTAGAGCAAAAGGAGCTCGAAAACA	1124
Qy	1304	CCGAACGCTGGGGCTCCTCAGGCCAATGGATGCCCTGGGTCTCTCCCTTCTTAGGACCTC	1363
Db	1125	CTGACCTCTGGGGCTCCTCAGCCAAATGGATGCCCTGGATCTCCCTTCTTAGGACCTC	1184
Qy	1364	TAGCAGCTCTAAATATTTACTCTCTTTTGGACCCGTATCTTTTAAACCTCCTTGTGTTAAGT	1423
Db	1185	TAGCAGCTATAATATTTGTTACTCTCTTTTGGACCCGTATCTTTTAAACCTCCTTGTGTTAAGT	1244
Qy	1424	TTGTCTCTTCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAATCGAACCCCA	1481
Db	1245	TTGTCTTTCCAGAAATCGAAGCAGTAAACTACAATTCGTTCTTCAATGGAGCCCCA	1302

RESULT 16

```

US-10-632-793-24
; Sequence 24, Application US/10632793
; Publication No. US20040048298A1
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucaia
APPLICANT: MALLET, Francois
APPLICANT: VOISSET, Cecile
TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
FILE REFERENCE: 110048
CURRENT APPLICATION NUMBER: US/10/632,793
CURRENT FILING DATE: 2003-08-04
PRIOR APPLICATION NUMBER: US/09/869,927
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: PCT/FR00/00144
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: FR 99/00888
PRIOR FILING DATE: 1999-01-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1

```

Query Match 75.4%; Score 1116.4; DB 17; Length 1948;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 71; Indels 0; Gaps 0

	Qy	Db
1	ATGGCGCTCCGTTATCATATCTTTCTCTTTACTGTCTCTTACCCCTCTCGCTCTCACT	60
719	ATGGCGCTCCGTTATCATATTTTCTTTTACTGTGTTCACCCCTCTTTACTCTCACT	778

QY 1141 CAAATCGAAGAGCTTTAGACTTCTAAGCGCAAGAGGGGGAACCTGTTTATTTTAA 1200
Db |||||||
Db 1859 CAAATCGAAGAGCTTTAGACTTCTAAGCGCTGAAGAGGGGGAACCTGTTTATTTA 1918
QY 1201 GGAGAGAAGAGCTGTTTATTTATGTTAAATCAA 1230
Db |||||||
Db 1919 GGGGAAGAAGCTGTTTATTTATGTTAAATCAA 1948

RESULT 17
US-09-864-761-21192
; Sequence 21192, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21192
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
; OTHER INFORMATION: NT HIT: AF208161.1, EVALUATE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: AUI38405.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUATE 2.00e-16
US-09-864-761-21192

Query Match 48.4%; Score 716.6; DB 9; Length 792;
Best Local Similarity 94.7%; Pred. No. 1.4e-209;
Matches 753; Conservative 0; Mismatches 39; Indels 3; Gaps 1;
QY 375 AGAAAAACAAGTAAGGAAGCAATCTCCCAACTGACCCCGGGGACATAGACCCCTAGCCCC 434
Db |||||||
Db 1 AGAAAAACAATAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGGCCCC 60
QY 435 CTACAAAGAGACTAGTCTCTCAAAATACATGAAGACCCCTCCGTAACCATCTCGCCTGGT 494
Db |||||||
Db 61 CTACAAAGAGACTAGATCTCTCAAAATACATGAAGACCCCTCCATACCATTAATGGCCTGGT 120
QY 495 GAGCCTATTATATACCAACCTCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCTACTAA 554
Db |||||||
Db 121 AGCCTATTATATACCAACCTCTCACTCGGCTCCATGAGGTCTCGGCCCAAAACCTACTAA 180
QY 555 CTGTTGGATGTGCTCTCCCTCCCTGCACTTCAGGGCCATACATTTCAATCCCTGTTCTGAACA 614
Db |||||||
Db 181 CTGTTGGATGTGCTCTCCCTCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240
QY 615 ATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGTTTGTAGTAGGACCTCTGTTTC 674
Db |||||||
Db 241 ATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGTTTGTAGTAGGACCTCTGTTTC 297
QY 675 CAATCTGGAAATAACCCATACCTCAAAACCTCACTGTGTAAATTTTAGCAATACTATAGA 734
Db |||||||
Db 298 CAATCTGGAAATAACCCATACCTCAAAACCTCACTGTGTAAATTTTAGCAATACTATAGA 357
QY 735 CACAACAGCTCCCAATGCAATGATGAGTGGGTAAACACCTCCACACGAATAGTGTGCTTACC 794
Db |||||||
Db 358 CACAGCCNACTCCCAATGCAATGATGAGTGGGTAACTCCTCCACACGAATAGTGTGCTTACC 417
QY 795 CTGAGGAATATTTTGTCTGTGTGATCTCAGGCTATCATTTGTAATGGCTCTTCAGA 854
Db |||||||
Db 418 CTGAGGAATATTTTGTCTGTGTGATCTCAGGCTATCATTTGTAATGGCTCTTCAGA 477
QY 855 ATCTATGTGCTCTCTCATTTCTAGTGGCCCTATGACCATCTACACTGAACAAGATTT 914
Db |||||||
Db 478 ATCTGTGTGCTCTCTCATTTCTAGTGGCCCTATGACCATCTACACTGAACAAGATTT 537
QY 915 ATACAATCATGTGCTACCTTAAGCCCAACAAACAAAGAGTACCCATTTCTTTTGTAT 974
Db |||||||
Db 538 ATACAATCATGTGCTACCTTAAGCCCAACAAACAAAGAGTACCCATTTCTTTTGTAT 597
QY 975 CAGAGCAGGAGTGTCTAGGAGAGCTAGGTACTGGCATTTGGCAGTATCAACCTCTACTCA 1034
Db |||||||
Db 598 TGGAGCAGGAGTGTCTAGGCGAGTACTACTGGCATTTGGCGGTATCAACCTCTACTCA 657
QY 1035 GTTCTACTCAAACTATCTCAAGAAATAAATGGTACATGCAACAGGTCACTGACTCCCT 1094
Db |||||||
Db 658 GTTCTACTCAAACTGTCTCAAGAACTAATGTTGACATGGAATGGGTGCGTGATACCT 717
QY 1095 GGTCACTCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTCAAAATCAAGAGC 1154
Db |||||||
Db 718 GGTCACTCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTCAAAATCAAGAGC 777
QY 1155 TTTAGACTTGTCTAAC 1169
Db |||||||
Db 778 TTTAGACTTGTCTAAC 792

RESULT 18
US-10-363-616-228
; Sequence 228, Application US/10363616

Fr1 Feb 25 16:26:32 2005

Publication No. US20040044181A1
 GENERAL INFORMATION:
 APPLICANT: Hyeq, Inc
 TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 FILE REFERENCE: 21272-113 (793)
 CURRENT APPLICATION NUMBER: US/10/363,616
 PRIOR FILING DATE: 2003-03-03
 PRIOR APPLICATION NUMBER: 09/654,935
 PRIOR FILING DATE: 2000-09-01
 NUMBER OF SEQ ID NOS: 490
 SEQ ID NO 228
 LENGTH: 1684
 TYPE: DNA
 ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (89)...(1684)
 US-10-363-616-228

Query Match 40.0%; Score 591.8; DB 17; Length 1684;
 Best Local Similarity 98.8%; Pred. No. 6.2e-171;
 Matches 596; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 879 AGTGGCCCTATGACCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCC 938
 DB 721 AGTGGCCCTATGACCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCC 780
 QY 939 CCACAAACAAAGAGTACCCATTCTTCTTTTGTATCAGAGCAGGAGTCTAGGAGACT 998
 DB 781 CCACAAACAAAGAGTACCCATTCTTCTTTTGTATCAGAGCAGGAGTCTAGGAGACT 840
 QY 999 AGTACTGGGAGTATGACATCAACACCTCTACTAGTCTTACTACAAATCTCTCAAGA 1058
 DB 841 AGTACTGGGAGTATGACATCAACACCTCTACTAGTCTTACTACAAATCTCTCAAGA 900
 QY 1059 AATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGCCTCAAGATCAACTTAA 1118
 DB 901 AATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGCCTCAAGATCAACTTAA 960
 QY 1119 CTCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAG 1178
 DB 961 CTCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAG 1020
 QY 1179 AGGGGAACCTGTTATTTTAGAGAGAACGCTGTTATTTATGTTAATCAATCCAGAAAT 1238
 DB 1021 AGGGGAACCTGTTATTTTAGAGAGAAATGCTGTTATTTATGTTAATCAATCCAGAAAT 1080
 QY 1239 TGTCACCTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCA 1298
 DB 1081 TGTCACCTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCA 1140
 QY 1299 AAACACCGAAGCGTGGGCTCTCTCAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGG 1358
 DB 1141 AAACACCGAAGCGTGGGCTCTCTCAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGG 1200
 QY 1359 ACCTTAGCAGCTCTAATATTTTACTCTCTTGGACCTCTGATCTTTAACCCTCTTGT 1418
 DB 1201 ACCTTAGCAGCTCTAATATTTTACTCTCTTGGACCTCTGATCTTTAACCCTCTTGT 1260
 QY 1419 TAAGTTTGTCTCTTCCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACC 1478
 DB 1261 TAAGTTTGTCTCTTCCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACC 1320
 QY 1479 CCA 1481
 DB 1321 CCA 1323

RESULT 19
 US-10-632-793-19
 ; Sequence 19, Application US/10632793
 ; Publication No. US200400482981
 ; GENERAL INFORMATION:
 ; APPLICANT: PARANHOS-BACCALA, Glaucia

APPLICANT: PARANHOS-BACCALA, Glaucia
 APPLICANT: MALLET, Francois
 APPLICANT: VOISSET, Cecile
 TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
 TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
 FILE REFERENCE: 110048
 CURRENT APPLICATION NUMBER: US/10/632,793
 CURRENT FILING DATE: 2003-08-04
 PRIOR FILING DATE: 2001-10-22
 PRIOR APPLICATION NUMBER: PCT/FR00/00144
 PRIOR FILING DATE: 2000-01-21
 PRIOR APPLICATION NUMBER: FR 99/00888
 PRIOR FILING DATE: 1999-01-21
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 19
 LENGTH: 591
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-632-793-19

Query Match 39.4%; Score 582.8; DB 17; Length 591;
 Best Local Similarity 99.7%; Pred. No. 2e-168;
 Matches 584; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 887 CTATGACCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCCCAACA 946
 DB 1 CCAATGGCCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCCCAACA 60
 QY 947 AAAGAGTACCCATTCTTCTTTTGTATCAGAGCAGGAGTCTAGGAGACTAGGTACTG 1006
 DB 61 AAAGAGTACCCATTCTTCTTTTGTATCAGAGCAGGAGTCTAGGAGACTAGGTACTG 120
 QY 1007 GCATGGGAGTATCAACACCTCTACTAGTCTTACTACAAATCTCTCAAGAAATAAATG 1066
 DB 121 GCATGGGAGTATCAACACCTCTACTAGTCTTACTACAAATCTCTCAAGAAATAAATG 180
 QY 1067 GTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGCCTTCAAGAGTCAACTTAACTCCCTAG 1126
 DB 181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGCCTTCAAGAGTCAACTTAACTCCCTAG 240
 QY 1127 CAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAA 1186
 DB 241 CAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAA 300
 QY 1187 CCGTGTATTTTAGGAGAGAAACGCTGTTATTTATTAATCAATCCAGAAATTTGCTACTG 1246
 DB 301 CCGTGTATTTTAGGAGAGAAACGCTGTTATTTATTAATCAATCCAGAAATTTGCTACTG 360
 QY 1247 AGAAGATTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCG 1306
 DB 361 AGAAGATTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCG 420
 QY 1307 AACGCTGGGCTCTCTCAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGGACCTCTAG 1366
 DB 421 AACGCTGGGCTCTCTCAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGGACCTCTAG 480
 QY 1367 CAGCTCTAATATTTTACTCTCTTTGGACCTCTGATCTTTAACCCTCTCTTTTAAAGTTTG 1426
 DB 481 CAGCTCTAATATTTTACTCTCTTTGGACCTCTGATCTTTAACCCTCTCTTTTAAAGTTTG 540
 QY 1427 TCTCTTCCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTCAAAAT 1472
 DB 541 TCTCTTCCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTCAAAAT 586

RESULT 20
 US-10-632-793-25
 ; Sequence 25, Application US/10632793
 ; Publication No. US200400482981
 ; GENERAL INFORMATION:
 ; APPLICANT: PARANHOS-BACCALA, Glaucia


```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322491
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(570)
; OTHER INFORMATION: n = A,T,C or G
;
US-10-027-632-322492

Query Match      25.5%; Score 377.8; DB 13; Length 570;
Best Local Similarity 90.7%; Pred. No. 3e-105;
Matches 411; Conservative 2; Mismatches 39; Indels 1; Gaps 1;

QY      1  ATGGCCCTCCCTTATCATACATCTTTCTCTTTACTGTTCTCTTACCCCTTTGCTCTCACT 60
DB      463 ATAGCCCTCCCTTATCATATTTTCTTTTACCAITCTCTTACCTGCTTCTCACT 404

QY      61  GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
DB      403 TCACCTCACTCCATGCTGTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCT-TGGAGA 345

QY      121 ACGGGCTTCTTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAAGGAAATCC 180
DB      344 ATGCAGCTTCTTGGAAATATTGATTTCCCATCATATAGGAGTTTATCTAAGGAAATCAC 285

QY      181 ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
DB      284 ACCTTCACCTGCCACACCCATATGCCCTTACCACTGCTATAAGCTTCCACTCTTTGCATG 225

QY      241 CATGCAAAATCACTATTATTTGGACAGGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT 300
DB      224 CATGCAAAATCACTAAATATTGTTTCATGGAAATGATTAATCTCTCTGGAGACTT 165

QY      301 GGAGCCACTGCTGTGGACTTACTTACCCCATACCACTACAGTAGTCTGTGATGGGGTGAAT 360
DB      164 GGAGCCACTGCTGTGTGGACTTACTTACCCCATACCACTACAGTAGTCTGTGATGGGGTGAAT 105

QY      361 CAAAGTCAGGCAAGAGAAAACAGTAAAGGAAGCAATCTCCCAACTGACCCGGGACAT 420
DB      104 CAAAGTCAGGCAAGAGAAAACAGTAAAGGAAGTAACTCTCCCAACTGACCCGGTACAT 45

QY      421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTC 453
DB      44  AGCACCCCTAGACCCCTACAAAGGACTAATCCCC 12

RESULT 24
US-10-027-632-322491/c
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322491
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(570)
; OTHER INFORMATION: n = A,T,C or G
;
US-10-027-632-322491

Query Match      25.5%; Score 377.8; DB 13; Length 570;
Best Local Similarity 90.7%; Pred. No. 3e-105;
Matches 411; Conservative 2; Mismatches 39; Indels 1; Gaps 1;

QY      1  ATGGCCCTCCCTTATCATACATCTTTCTCTTTACTGTTCTCTTACCCCTTTGCTCTCACT 60
DB      463 ATAGCCCTCCCTTATCATATTTTCTTTTACCAITCTCTTACCTGCTTCTCACT 404

QY      61  GCACCCCTCCATGCTGTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
DB      403 TCACCTCACTCCATGCTGTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCT-TGGAGA 345

QY      121 ACGGGCTTCTTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAATCC 180
DB      344 ATGCAGCTTCTTGGAAATATTGATTTCCCATCATATAGGAGTTTATCTAAGGAAATCAC 285

QY      181 ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
DB      284 ACCTTCACCTGCCACACCCATATGCCCTTACCACTGCTATAAGCTTCCACTCTTTGCATG 225

QY      241 CATGCAAAATCACTATTATTTGGACAGGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT 300
DB      224 CATGCAAAATCACTAAATATTGTTTCATGGAAATGATTAATCTCTCTGGAGACTT 165

QY      301 GGAGCCACTGCTGTGGACTTACTTACCCCATACCACTACAGTAGTCTGTGATGGGGTGAAT 360
DB      164 GGAGCCACTGCTGTGTGGACTTACTTACCCCATACCACTACAGTAGTCTGTGATGGGGTGAAT 105

QY      361 CAAAGTCAGGCAAGAGAAAACAGTAAAGGAAGCAATCTCCCAACTGACCCGGGACAT 420
DB      104 CAAAGTCAGGCAAGAGAAAACAGTAAAGGAAGTAACTCTCCCAACTGACCCGGTACAT 45

QY      421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTC 453
DB      44  AGCACCCCTAGACCCCTACAAAGGACTAATCCCC 12

RESULT 23
US-10-027-632-322492/c
; Sequence 322492, Application US/10027632
```



```

RESULT 26
US-10-029-386-4312/c
; Sequence 4312, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4312
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021619.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: EST HUMAN HIT: BE734284.1, EVALUATION 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUATION 4.00e-26
; OTHER INFORMATION: NT HIT: AF208161.1, EVALUATION 0.00e+00
US-10-029-386-4312

Query Match 24.6%; Score 364; DB 16; Length 521;
Best Local Similarity 91.7%; Pred. No. 5.2e-101;
Matches 385; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1062 AAATGTGACATGGACAGGTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAATC 1121
DB 521 AAATGTGACATGGATGGTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAATC 462
QY 1122 CCTAGCAGCAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTAAACGCCAAAGAGG 1181
DB 461 CTTAGCAGCAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTAAACGCCAAAGAGG 402
QY 1182 GGGAACTGTTATTTTATAGGAGAGAGCGCTGTATTATTTATTAATCAATCCAGAAATGT 1241
DB 401 GGGAACTGTTATTTTATAGGAGAGAGCGCTGTATTATTTATTAATCAATCCAGAAATGT 342
QY 1242 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGAGAGGCTTCAAAA 1301
DB 341 CACCGAAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGAGAGGCTTCAAAA 282
QY 1302 CACCGAACGCTGGGCGCTCTCAGCAATGGATGGCTGGCTTCTCCCTTCTTAGGACC 1361
DB 281 CACCGAACGCTGGGCGCTCTCAGCAATGGATGGCTGGCTTCTCCCTTCTTAGGACC 222
QY 1362 TCTAGCAGCTCTAATATTTTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTTAA 1421
DB 221 TCTAGCAGCTCTAATATTTTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTTAA 162
QY 1422 GTTTGTCTCTCAGAAATTCGAGCTGTAAGCTTACAGATGGCTTCAAAATGGAACCCCA 1481
DB 161 GTTTGTCTCTCAGAAATTCGAGCTGTAAGCTTACAGATGGCTTCAAAATGGAACCCCA 102

RESULT 27
US-10-027-632-322574
; Sequence 322574, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

```

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322574
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(551)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-322574

Query Match 23.8%; Score 353; DB 13; Length 551;
Best Local Similarity 92.0%; Pred. No. 1.3e-97;
Matches 412; Conservative 2; Mismatches 30; Indels 4; Gaps 4;

QY 1 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTTTTACCCCTTTTGCCTCTCACT 60
DB 102 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTTTTACCCCTTTTGCCTCTCACT 161
QY 61 GCACCCCTCCATCTGCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 162 GCACCCCTCCATCTGCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGRAGA 221
QY 121 ACGGCGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 180
DB 222 ATGGCGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 280
QY 181 ACCTTCACTGCCACACCCATATGCCCGGACCTGCTATACTGTCACCTCTTTGCATG 240
DB 281 ACCTTCACTGCCACACCCATATGCCCGGACCTGCTATACTGTCACCTCTTTGCATG 340
QY 241 CATGCAATATCTATTATTGGACAGGGAATGATTAATCTTAGTTGCTTGGAGACTT 300
DB 341 CATGCAATATCTATTATTGGACAGGGAATGATTAATCTTAGTTGCTTGGAGACTT 400
QY 301 GGAGCCACTGCTGTGGACTT-ACCTCACCCATACCAAGTATGTCTGATGGGGTGGAT 359
DB 401 GGAGCCACTGCTGTGGACTTAACTTCACTCATACCGGTATGTCTGATGGGGTGGAGT 460
QY 360 TCAAGGTGAGCAAGAGAAAACAAAGTAAGGAGCAATCTCCCAACTGACCCGGGACA 419
DB 461 TC-AGATGAGGCAAGAG-AAAACACGTANAGGAAGTAAATCTCCANCTGACNGTGTACA 518
QY 420 TAGCACCCCTAGCCCTACAAAGGACTA 447
DB 519 TAGCACCCCTAACCCCTACAAAGGACTA 546

RESULT 28
US-10-027-632-322575
; Sequence 322575, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

```


;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 322574
;; LENGTH: 551
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(551)
;; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-322574

Query Match 23.8%; Score 353; DB 17; Length 551;
Best Local Similarity 92.0%; Pred. No. 1.3e-97;
Matches 412; Conservative 2; Mismatches 30; Indels 4; Gaps 4;

QY 1 ATGGCCCTCCCTTATCATATCTTTCTTCTTACTGTTCTTACCCCTTTGGCTCTCACT 60
DB |||||
QY 102 ATGGCCCTCCCTTATCATATTTTCTTCTTACTGTTCTTACCCCTTTCACTCTCACT 161
DB |||||
QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB |||||
QY 162 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAATAGTTTCTATGAGA 221
DB |||||
QY 121 ACAGCGCTTCTGGAATATTGATGCCCATCATATAGAGTTTATCTAAGGAACTCC 180
DB |||||
QY 222 ATGGCGCTTCTGGAATATTGATGCCCATCATATAGAGTTTATCTAAGGAACTCC 280
DB |||||
QY 181 ACCTTCACTGCCACACCATATGCCCCGCACTGCTATAAATCTGCACTCTTTGCAATG 240
DB |||||
QY 281 ACCTTCACTGCCACACCATATGCCCCGCACTGCTATAAATCTGCACTCTTTGCAATG 340
DB |||||
QY 241 CATGCAAAATCTCAATATTGGACAGGAAATGATTAATCTAGTTGCTCTGAGGACTT 300
DB |||||
QY 341 CATGCAAAATCTCAATATTGGACAGGAAATGATTAATCTAGTTGCTCTGAGGACTT 400
DB |||||
QY 301 GGAGCCACTGCTGTTGGACTT-ACCTCACCATACAGTATGCTGATGGGGTGAAT 359
DB |||||
QY 401 GGAGCCACTGCTGTTGGACTTAACTTCACTATACCGGTATGCTGATGGGGTGGAGT 460
DB |||||
QY 360 TCAGGTCAGGCAAGAGAAACAACTAAAGGAACTCTCCAACTGACCCGGGGA 419
DB |||||
QY 461 TC-AGATGAGGCAAGAG-AAAACACGTANAGGAAATTAATCTCANTGACNGTGTACA 518
DB |||||
QY 420 TAGCACCCCTAGCCCTTACAAAGGACTA 447
DB |||||
QY 519 TAGCACCCCTAACCCCTACAAAGGACTA 546
DB |||||

RESULT 30
US-10-027-632-322575
; Sequence 322575, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: 108827.129

;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 322575
;; LENGTH: 551
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(551)
;; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-322575

Query Match 23.8%; Score 353; DB 13; Length 551;
Best Local Similarity 92.0%; Pred. No. 1.3e-97;
Matches 412; Conservative 2; Mismatches 30; Indels 4; Gaps 4;

QY 1 ATGGCCCTCCCTTATCATATCTTTCTTCTTACTGTTCTTACCCCTTTGGCTCTCACT 60
DB |||||
QY 102 ATGGCCCTCCCTTATCATATTTTCTTCTTACTGTTCTTACCCCTTTCACTCTCACT 161
DB |||||
QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB |||||
QY 162 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAATAGTTTCTATGAGA 221
DB |||||
QY 121 ACAGCGCTTCTGGAATATTGATGCCCATCATATAGAGTTTATCTAAGGAACTCC 180
DB |||||
QY 222 ATGGCGCTTCTGGAATATTGATGCCCATCATATAGAGTTTATCTAAGGAACTCC 280
DB |||||
QY 181 ACCTTCACTGCCACACCATATGCCCCGCACTGCTATAAATCTGCACTCTTTGCAATG 240
DB |||||
QY 281 ACCTTCACTGCCACACCATATGCCCCGCACTGCTATAAATCTGCACTCTTTGCAATG 340
DB |||||
QY 241 CATGCAAAATCTCAATATTGGACAGGAAATGATTAATCTAGTTGCTCTGAGGACTT 300
DB |||||
QY 341 CATGCAAAATCTCAATATTGGACAGGAAATGATTAATCTAGTTGCTCTGAGGACTT 400
DB |||||
QY 301 GGAGCCACTGCTGTTGGACTT-ACCTCACCATACAGTATGCTGATGGGGTGAAT 359
DB |||||
QY 401 GGAGCCACTGCTGTTGGACTTAACTTCACTATACCGGTATGCTGATGGGGTGGAGT 460
DB |||||
QY 360 TCAGGTCAGGCAAGAGAAACAACTAAAGGAACTCTCCAACTGACCCGGGGA 419
DB |||||
QY 461 TC-AGATGAGGCAAGAG-AAAACACGTANAGGAAATTAATCTCANTGACNGTGTACA 518
DB |||||
QY 420 TAGCACCCCTAGCCCTTACAAAGGACTA 447
DB |||||
QY 519 TAGCACCCCTAACCCCTACAAAGGACTA 546
DB |||||

RESULT 29
US-10-027-632-322574
; Sequence 322574, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632

```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32575
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(551)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-32575

Query Match      23.8%; Score 353; DB 17; Length 551;
Best Local Similarity 92.0%; Pred. No. 1.3e-97;
Matches 412; Conservative 2; Mismatches 30; Indels 4; Gaps 4;

QY 1 ATGGCCCTCCCTTATCATACTTTCTTTCTTACTGTCTCTTACCCCTTCCTCTCACT 60
Db 102 ATGGCCCTCCCTTATCATATATTTTCTTTCTTACTGTCTCTTACCCCTTCCTCTCACT 161

QY 61 GCACCCCTCCATGCTGTGTACAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 162 GCACCCCTCCATGCTGTGTACGACAGTAGTCCCTTACCAATAGTTTCTATGRAGA 221

QY 121 ACCTGCGCTTCTTGGAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 222 ATGCGGCTTCTTGGAATATTGATGCCCCCATCATCAWATAGGAGTTTATCTAAGGGAAC-CC 280

QY 181 ACCTTCACTGCCACACCCCATATGCCCCGCACTGTATTAACCTCTGCACTCTTTGCATG 240
Db 281 ACCTTCACTGCCACACCCCATATGCCCCCAAAATGCTGTAACTCTGCCACTCTTTGCATG 340

QY 241 CATGCAAACTACTCATTTATGGACAGGAAATGATTAACTCTAGTTGTCTGAGGACTT 300
Db 341 CATGCAAACTACTCATTTATGGACAGGAAATGATTAACTCTAGTTGTCTTGGAGACTT 400

QY 301 GGAGCCACTGTCTGTGGACTT-ACCTCAACCCATACAGTATGTCTGTATGGGGTGGAAAT 359
Db 401 GGAGCCACTGTCTGTGGACTTAACTCACTATACCGGTATGTCTGTATGGGGTGGAGT 460

QY 360 TCAAGTCAGGCAAGAGAAAAAACAAGTAAGGAAGCAATCTCCCACTGACCCGGGACA 419
Db 461 TC-AGATGAGGCAAGAG-AAAACACGTANAGGAAGTAAATCTCCACTGACNGTGGTACA 518

QY 420 TAGCACCCCTAGCCCTTACAAAGACTA 447
Db 519 TAGCACCCCTAACCCCTTACAAAGACTA 546

RESULT 31
US-10-292-798-1393/c
; Sequence 1393, Application US/10292798
; Publication No. US200302583A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
```

```

; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1393
; LENGTH: 822900
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(822900)
; NAME/KEY: CDS
; LOCATION: (201)..(1068)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (140545)..(140693)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (261786)..(261845)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (273663)..(273702)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (347633)..(347711)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482589)..(482596)
; NAME/KEY: CDS
; LOCATION: (534176)..(534210)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (822485)..(822700)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (4848)..(4947)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (4966)..(4966)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (17785)..(17884)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (61159)..(61258)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (67605)..(67704)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (74625)..(74724)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (85854)..(85953)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
```

LOCATION: (367573)..(367573)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (367588)..(367588)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (367595)..(367595)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (367613)..(367614)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (367985)..(367985)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (367993)..(367993)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (367995)..(367996)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (367998)..(367998)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (370273)..(370277)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (370279)..(370279)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (370281)..(370282)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (370285)..(370287)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (370289)..(370291)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (385643)..(385742)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (409961)..(410060)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (410096)..(410096)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (417384)..(417483)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (724960)..(725059)	OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:	
NAME/KEY: modified base	
LOCATION: (726106)..(726205)	OTHER INFORMATION: a, t, c, g, unknown or other

QY 1073 TGAACAGGTCACTGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGCAG 1132
DB 56723 TGAATGAAATGCCAACTCCCTAGTGACCTTACAAAGCCAGCTTAATTCTCTAGCTGCAG 56664
QY 1133 TAGTCTCTTCAAAATCGAAGAGCTTTAGACTTTGCTTAACCGCCCAAAAGAGGGGGAACCTGTT 1192
DB 56663 TAGTCTCTTCAAAACCGGAGAGCCCTAGACTTATTAAACAGCTAAATAAGGAGGAACCTGCC 56604
QY 1193 TATTTTGGAGAGAACGCTGTTATATGTTAATCAATCCAGATTGTCTACTGAGAAAG 1252
DB 56603 TCTTCTTAGGAAAAGAAATGTTGCTATTTTGGTTTAAACAGCTCAGGAATCATTACTGAAAAG 56544
QY 1253 TTAAGAAATTCGAGATCGAATACAAATGATAGAGAGAGAGCTTCAAAACACCGAACGCT 1312
DB 56543 TCAAGAAATAAGAGAACAAATAGAAAGTAGAAAAGAGAGCTTGAACACTCAGGGCGCT 56484
QY 1313 GGGGCGCTCTCAGCAATGGATGCCCTCGGTTCCTCCCTTCTTAGGACCTCTAGCAGCTC 1372
DB 56483 TGAATATGTTTAAACCAATGGATACCTCGGCTCCTCCCTTCTAGGCGCTGCGACAGC-- 56426
QY 1373 TAATATGTTACTCTCTTTGGACCGCTGATCTTTAACTCCTCTTGTAAAGTTTGTCTTT 1432
DB 56425 ----CATCTACTCGTTTTTGGGCGCTTGCAITTTTAACTCCTCTGTCAAAITTTGTTCT 56370
QY 1433 CCAGAAATCAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
DB 56369 -CAGGATCAGGCGCATCAAGGTACAAATGGTCTTACAAATGGAACCTCA 56322

RESULT 32
US-08-979-847-108
; Sequence 108, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1329 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-979-847-108
Query Match 20.4%; Score 302; DB 8; Length 1329;
Best Local Similarity 92.7%; Pred. No. 1.1e-81;
Matches 317; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1140 TCAAAATCGAAGAGCTTTAGACTTGTACCTGCTAACCGCCCAAAAGAGGGGGAACCTGTTTATTTT 1199
DB 1 TCAAAATCGAAGAGCTTTAGACTTGTACCTGCTAACCGCCCAAAAGAGGGGGAACCTGTTTATTTT 60
QY 1200 AGGAGAAAGCGCTGTTTATTATGTTAATCAATCCAGAAATGTCTCACTGAGAAAGTTAAAGA 1259
DB 61 AGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCATYACTGGAAGTTAAAGA 120
QY 1260 AATTTCGAGATCGAATAACAATGTAGAGCAGAGAGCTTTCAAAACACCGAACCTCGGGCCT 1319
DB 121 AATTTCGAGATCGAATAACAATGTAGAGCAGAGAGCTTTCAAAACACCTCGGGCCT 180
QY 1320 CCTCAGCAATGGAATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAAATATT 1379
DB 181 CCTCAGCAATGGAATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAAATATT 240
QY 1380 GTTACTCTCTTTGGACCGCTGTATCTTTAACTCCTCTTAAAGTTTGTCTCTCCAGAAAT 1439
DB 241 TTTACTCTCTTTGGACCGCTGTATCTTTCACTTCTTGTAAAGTTTGTCTCTCCAGAAAT 300
QY 1440 TGAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
DB 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCCA 342

RESULT 33
US-10-114-104-108
; Sequence 108, Application US/10114104
; Publication No. US20030198647A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,104
; FILING DATE: 03-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,847
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:

[illegible]


```

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86541
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-86541

Query Match      15.5%; Score 229.6; DB 13; Length 619;
Best Local Similarity 86.3%; Pred. No. 1.5e-59;
Matches 253; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 317 GGACTTACTTACCCATACCCAGTATGTCGTATGGGGTGGAAATTCAGGTCAGGCAAGAG 376
Db 619 GAATTCACCTTACCCATACCCAGTATGTCGTATGGGGTGGAGTTCAAGATCAGGCAAGAG 560

QY 377 AAAACAAGTAAGGAAGCAATCTCCCAACTGACCCGGGGACATAGCACCCTAGCCCT 436
Db 559 AAAACAATGTAAGGAAGTCACTCCCAACTGACCTTGGGTACATTAAGCACCCTAGCCCT 500

QY 437 ACAAGGACTAGTCTCTCAAACTACATGAACCCCTCCGTACCATCTCGCTGTGTA 496
Db 499 GCAAGGACTAGTCTCTCAAACTACATGAACCCCTCCGTACCATCTCGCTGTGTA 440

QY 497 GCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCTTAACT 556
Db 439 GCCTATTATTAACCAACCTCACTCGGCTCCATNGAAGCCCTCAGAGAAAACCTTAACT 380

QY 557 GTTGGATGTGCTCCCTCGCTGACCTTCAGGCCATACATTTCAATCCCTGTTCCT 609
Db 379 GTTGGATGTGCTCCCTCGCTGACCTTCAGGCCATACATTTCAATCCCTGTATCT 327

```

```

RESULT 38
US-10-027-632-86541/c
; Sequence 86541, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86541
; LENGTH: 619

```

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-86541

Query Match      15.5%; Score 229.6; DB 17; Length 619;
Best Local Similarity 86.3%; Pred. No. 1.5e-59;
Matches 253; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 317 GGACTTACTTACCCATACCCAGTATGTCGTATGGGGTGGAAATTCAGGTCAGGCAAGAG 376
Db 619 GAATTCACCTTACCCATACCCAGTATGTCGTATGGGGTGGAGTTCAAGATCAGGCAAGAG 560

QY 377 AAAACAAGTAAGGAAGCAATCTCCCAACTGACCCGGGGACATAGCACCCTAGCCCT 436
Db 559 AAAACAATGTAAGGAAGTCACTCCCAACTGACCTTGGGTACATTAAGCACCCTAGCCCT 500

QY 437 ACAAGGACTAGTCTCTCAAACTACATGAACCCCTCCGTACCATCTCGCTGTGTA 496
Db 499 GCAAGGACTAGTCTCTCAAACTACATGAACCCCTCCGTACCATCTCGCTGTGTA 440

QY 497 GCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCTTAACT 556
Db 439 GCCTATTATTAACCAACCTCACTCGGCTCCATNGAAGCCCTCAGAGAAAACCTTAACT 380

QY 557 GTTGGATGTGCTCCCTCGCTGACCTTCAGGCCATACATTTCAATCCCTGTTCCT 609
Db 379 GTTGGATGTGCTCCCTCGCTGACCTTCAGGCCATACATTTCAATCCCTGTATCT 327

```

```

RESULT 39
US-09-864-761-7501/c
; Sequence 7501, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7501
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL139245.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; US-09-864-761-7501

Query Match      14.8%; Score 219.4; DB 9; Length 559;
Best Local Similarity 91.7%; Pred. No. 2e-56;
Matches 232; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      1229 AATCCAGAAATGTCACAGAGAGTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAG 1288
Db
QY      1289 AGGAGCTTCAAAACACCGAACCGCTGGGCGCTCTCAGCAATGGATGCCCTGGTCTCC 1348
Db
QY      1349 CTTCTTTAGGACCTCTAGCAGCTTAATATTTGTTACTCTTTGGACCCCTGTATCTTTA 1408
Db
QY      1409 ACTCTCTGTTAAGTTGTCTCTCCAGAAATTCAGAGCTGTAAGCTAGAGTGGCTTAC 1468
Db
QY      1469 AATGGAAACCCCA 1481
Db      319 AATGGAGCCCCA 307

RESULT 40
US-10-029-386-20259
; Sequence 20259, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20259
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; OTHER INFORMATION: MAP TO AP001697.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.91
; OTHER INFORMATION: NT HIT: AF08161.1, EVALUE 1.00e-108
; OTHER INFORMATION: SWISSPROT HIT: P28808, EVALUE 6.00e-12
; OTHER INFORMATION: EST_HUMAN HIT: BE734284.1, EVALUE 1.00e-108
; US-10-029-386-20259

Query Match      14.4%; Score 213; DB 16; Length 494;
Best Local Similarity 85.6%; Pred. No. 1.7e-54;
Matches 261; Conservative 0; Mismatches 40; Indels 4; Gaps 2;

QY      1106 AAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGC 1165
Db      193 AAGGACTTCAAAAGCGGTGTATGTCAGTAGTCCTTCAAAATCGAAGAGCTTTAGAAATTGC 252
QY      1166 TAACCGCCAAAAGAGGGGAAACCTGTTTATTTTAGGAGAAGAACGCTGTTATTATGTTA 1225
Db      253 TAATCACTGAGAGAGGGGAAACGTTTTTATTTTAGGGGAAGAAATGCTGTTATTATGTTA 312
QY      1226 ATCAATCCAGAAATGTCACCTGAGAGAAAGTTAAAGAAATTCGAGATCGAATCAAAATGTAGAG 1285
Db      313 ATCAATTCGGAAATCATCACCAGAAAGTTAAAGAAATTCAGATCGAATCAACGTAGAA 372
QY      1286 CAGAGGAGCTT-CAAAACACCGAACGCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGGTT 1344
Db      373 CAGAGGAGCTTAAAAAACACTGGACCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGATT 432
QY      1345 CTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTGTTACTCTCTTTGGACCCCTGTATC 1404
Db      433 CTCCCTTTCTTAGGACCTCTAGCAGCT---ATATTTCTACTCTCTTTGGACCCCTGTATC 489
QY      1405 TTTAA 1409
Db      490 TTTAA 494

Search completed: February 21, 2005, 19:18:14
Job time : 827.526 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:52:54 ; Search time 4605.5 Seconds
(without alignments)
12240.400 Million cell updates/sec

Title: US-09-319-156b-9
Perfect score: 1481
Sequence: 1 atggccctccctatcatcac.....gtcttacaatgaagcccca 1481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST.*
- 1: gb_est1.*
 - 2: gb_est2.*
 - 3: gb_est3.*
 - 4: gb_est4.*
 - 5: gb_est5.*
 - 6: gb_est6.*
 - 7: gb_est7.*
 - 8: gb_est8.*
 - 9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1333.8	90.1	2500	3 CR622175	full-length cDNA clone CS0D1051YM13 of Placentia Cot 25-normalized
2	1333.8	90.1	2716	3 CR613169	full-length cDNA libraries and normalization
3	1333.8	90.1	2748	3 CR605851	full-length cDNA libraries and normalization
4	1333.8	90.1	2748	3 CR625046	full-length cDNA libraries and normalization
5	1333.8	90.1	2749	3 CR617248	full-length cDNA libraries and normalization
6	839	56.7	935	5 BX391741	full-length cDNA libraries and normalization
7	813.6	54.9	955	5 BX347952	full-length cDNA libraries and normalization
8	752.2	50.8	931	5 BX389656	full-length cDNA libraries and normalization
9	747.2	50.5	872	5 BX326647	full-length cDNA libraries and normalization
10	741	50.0	903	5 BX408752	full-length cDNA libraries and normalization
11	737.2	49.8	883	5 BX408733	full-length cDNA libraries and normalization
12	728.4	49.2	877	5 BX347619	full-length cDNA libraries and normalization
13	714	48.2	836	5 BX388766	full-length cDNA libraries and normalization
14	709.8	47.9	921	5 BX408734	full-length cDNA libraries and normalization
15	705.8	47.7	844	5 BX347111	full-length cDNA libraries and normalization
16	703.2	47.5	828	5 BX367907	full-length cDNA libraries and normalization
17	698.6	47.2	846	5 BX368078	full-length cDNA libraries and normalization
18	696.2	47.0	995	5 BX389657	full-length cDNA libraries and normalization
19	695	46.9	924	5 BX347314	full-length cDNA libraries and normalization
20	647	43.7	723	5 AUI138405	full-length cDNA libraries and normalization
21	629.4	42.5	631	8 AQ261133	full-length cDNA libraries and normalization
22	628.6	42.4	702	1 AUI138097	full-length cDNA libraries and normalization
23	623.8	42.1	924	5 BX409328	full-length cDNA libraries and normalization
24	609.8	41.2	714	2 BE734284	full-length cDNA libraries and normalization

C 25	580.4	39.2	658	5	BX430050	BX430050
C 26	555.8	37.5	792	5	BX409304	BX409304
C 27	537.6	36.3	814	5	BX430055	BX430055
C 28	536.4	36.2	586	4	BE019603	BE019603
C 29	535.2	36.1	790	4	BG572445	BG572445
C 30	478.6	32.3	760	5	BX450495	BX450495
C 31	475.4	32.1	554	7	CR735592	CR735592
C 32	444	30.0	1071	5	BX365066	BX365066
C 33	421.6	28.5	944	5	BX368759	BX368759
C 34	416.4	28.1	890	5	BX326300	BX326300
C 35	412.2	27.8	998	5	BX337769	BX337769
C 36	412.2	27.8	1058	5	BX378303	BX378303
C 37	411.4	27.8	905	5	BX429316	BX429316
C 38	406.4	27.4	564	5	BX388917	BX388917
C 39	402	27.1	1019	5	BX439636	BX439636
C 40	401.4	27.1	484	5	BX409035	BX409035
C 41	399.6	27.0	494	5	BQ365207	BQ365207
C 42	390.2	26.3	427	1	AA776439	AA776439
C 43	389.8	26.3	718	1	AL601847	AL601847
C 44	373.2	25.2	966	5	BX380176	BX380176
C 45	371.4	25.1	586	8	AQ776289	AQ776289

ALIGNMENTS

RESULT 1
CR622175
LOCUS
DEFINITION
full-length cDNA clone CS0D1051YM13 of Placentia Cot 25-normalized
of Homo sapiens (human)
ACCESSION
CR622175
VERSION
CR622175.1 GI:50502982
KEYWORDS
HTC; CNSLT_cDNA
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 2500)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 2500)
REFERENCE
2 (bases 1 to 2500)
AUTHORS
Direct Submission
TITLE
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the PCVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..2500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YM13"
/tissue type="Placentia Cot 25-normalized"
/plasmid="PCVSPORT_6"

ORIGIN
Query Match 90.1%; Score 1333.8; DB 3; Length 2500;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGCCCTCCCTATCATATTTCTTTACTGTTCTTACCCCTTTCCTACT 60
|||||
Db 530 ATGCCCTCCCTATCATATTTCTTTACTGTTCTTACCCCTTTCCTACT 589
|||||

```
QY 61 GCACCCCTCCATGCTGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 590 GCACCCCTCCATGCTGCTGTATGACACAGTAGCTCCCTTACCAAGAGTTTCTATGGAGA 649
QY 121 AGCGGCTTCTGGAAATATTTGATGCCCATCATATAGAGTTTATCTAAGGAACTCC 180
Db 650 ATGCGAGCGTCCGGAATATTTGATGCCCATCGTATAGAGTCTTCTAAGGAACTCC 709
QY 181 ACCTTCACTGCCACACCCATATGCGCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
Db 710 ACCTTCACTGCCACACCCATATGCGCGCAACTGCTATAACTCTGCCACTCTTTGCATG 769
QY 241 CATGAAATACATATTTGGACAGGAAATGATTAATCTAGTGTCTCGGAGGACTT 300
Db 770 CATGAAATACATATTTGGACAGGAAATGATTAATCTAGTGTCTCGGAGGACTT 829
QY 301 GGAGCCACTGCTGTGGACTTACTTCACCCATACAGTATGCTCTGATGGGGTGAAT 360
Db 830 GGAGTCACTGCTGTGGACTTACTTCACCCATACAGTATGCTCTGATGGGGTGAAT 889
QY 361 CAAAGTCAAGCAAGAAACAAAGTAAAGGAAGCAATCTCCAACTGACCGGGGACAT 420
Db 890 CAAAGTCAAGCAAGAAACAAAGTAAAGGAAGTAACTCCCAACTCACCGGGTACAT 949
QY 421 AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAACTACATGAACCTCCGTACC 480
Db 950 GGCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACATGAACCTCCGTACC 1009
QY 481 CATACTCGCTGTGAGCTTATTAATACACCTCACCTCGGCTCCATGAGTCTCAGCC 540
Db 1010 CATACTCGCTGTGAGCTTATTAATACACCTCACCTCGGCTCCATGAGTCTCGGCC 1069
QY 541 CAAACCCCTACTAATCTGTGGTGTGCTCCCTGCACTTCAGGCTACATATTCATC 600
Db 1070 CAAACCCCTACTAATCTGTGGTGTGCTCCCTGCACTTCAGGCTACATATTCATC 1129
QY 601 CTGTGCTGAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
Db 1130 CTGTGCTGAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 1189
QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTT 720
Db 1190 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTT 1249
QY 721 AGCAATCTATAGACACAACTCCCAATGCAATCAGTGGGTAAACACTCCACACGA 780
Db 1250 AGCAATCTATACATACACAACTCCCAATGCAATCAGTGGGTAACTCCTCCACACAA 1309
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTGCTACCTCAGCCTATCATTTTG 840
Db 1310 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTGCTACCTCAGCCTATCATTTTG 1369
QY 841 AATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTATGTGCCCCCTATGACCACTAC 900
Db 1370 AATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTATGTGCCCCCTATGACCACTAC 1429
QY 901 ACTGAAACAAGATTATACAAATCATGTGCTACCTTAAGCCCAACAAACAAAGATACCAT 960
Db 1430 ACTGAAACAAGATTATACAAATCATGTGCTACCTTAAGCCCAACAAACAAAGATACCAT 1489
QY 961 CTTCCCTTTTGTATCAGACAGAGTGCTAGGCAGACTAGGTACTGGCATTTGGCATGATC 1020
Db 1490 CTTCCCTTTTGTATAGGACAGAGTGCTAGGTGCTAGGTACTGGCATTTGGCGGTATC 1549
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGAACAG 1080
Db 1550 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGGACATGGAACG 1609
QY 1081 GTCACTGACTCCCTGGTCACTTGTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1140
Db 1610 GTGCCGACTCCCTGGTCACTTGTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1669
```

```
QY 1141 CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTCTTTATTTTA 1200
Db 1670 CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTCTTTATTTTA 1729
QY 1201 GGAGAAGAACGCTGTATTATTAATCAATCAAGATTTGTCTACTGAGAAAGTTAAAGAA 1260
Db 1730 GGGAGAAGATGCTGTATTATTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1789
QY 1261 ATTGAGATCGAATACAACTAGACAGAGAGCTTCAAAACACCGAACCTCGGGGCTC 1320
Db 1790 ATTGAGATCGAATACAACTAGACAGAGAGCTTCAAAACACTGGACCTCGGGGCTC 1849
QY 1321 CTGAGCCAATGGATGCGCTGGGTTCTCCCTCTTCTAGGACCTTAGCAGCTCTAAATTG 1380
Db 1850 CTGAGCCAATGGATGCGCTGGATTTCTCCCTCTTCTAGGACCTTAGCAGCTCTAAATTG 1909
QY 1381 TTACTCTCTTTGAGCCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGAAAT 1440
Db 1910 CTACTCTCTTTGAGCCCTGTATCTTTAACTCTCTTGTAACTCTCTCTCCAGAAATC 1969
QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 1970 GAAGCTGTAAGCTACAAATGGAACCCCAAGATGCAATGCCAA 2010

RESULT 2
CR613169 2716 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODE013Y120 of Placenta of Homo sapiens
DEFINITION
ACCESSION CR613169
VERSION CR613169.1 GI:50493976
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 2716)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sedref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dr) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..2716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODE013Y120"
/tissue type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 90.1%; Score 1333.8; DB 3; Length 2716;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTCTTCTTACCCCTTTCGCTCTCACT 60
Db 773 ATGGCCCTCCCTTATCATATTTTCTTCTTACTGTCTTCTTACCCCTTTCCTCACT 832
```

QY 61 GCACCCCTCCATGCTGTGTACCAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 833 GCACCCCTCCATGCTGTGTGTACCAAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 892
QY 121 AGCGGCTTCTGGAATAATTTGATGCCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 893 ATGACGGTCCCGGAATAATTTGATGCCCCATCATATAGGAGTTTCTAAGGGAACCTCC 952
QY 181 ACCTTCACTGCCACACCCATATGCCCCCACTGCTATAAATCTCCCACTCTTTTCATG 240
Db 953 ACCTTCACTGCCACACCCATATGCCCCCACTGCTATAAATCTCCCACTCTTTTCATG 1012
QY 241 CATGCAAACTACTCATTTATGGAAGGGAATAATGATTAATCTAGTTGTCCTGAGGACTT 300
Db 1013 CATGCAAACTACTCATTTATGGAAGGGAATAATGATTAATCTAGTTGTCCTGAGGACTT 1072
QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCCATACAGATATGCTGATGGGGTGAAT 360
Db 1073 GGAGTCACTGTCTGTGGACTTACTTCAACCCATACAGATATGCTGATGGGGTGAAT 1132
QY 361 CAAGGTCAGCAAGAGAAAACAAGTAAGAGGACATCTCCCACTGACCCCGGACAT 420
Db 1133 CAAGATCAGCAAGAGAAAACAAGTAAGAGGATTAATCTCCCACTGACCCCGGATCAT 1192
QY 421 AGCACCCCTAGCCCTACAAAGACTAGTTCTCTCAAACTACATGAAACCTCCGTACC 480
Db 1193 GGACCTCTAGCCCTACAAAGACTAGATCTCTCAAACTACATGAAACCTCCGTACC 1252
QY 481 CATACTCGCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCACGCC 540
Db 1253 CATACTCGCTGTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCACGCC 1312
QY 541 CAAAACCTCTACTACTGTGAGTGTGCTCCCTCGACTTCAGGCCATACATTTCAATC 600
Db 1313 CAAAACCTCTACTACTGTGAGTGTGCTCCCTCGAACTTCAGGCCATATGTTTCAATC 1372
QY 601 CTGTTCTCTGAACAATGGAACAACCTTCAGACACAGAAATAACACCACTTCGTTTGTAGTA 660
Db 1373 CTGTTCTCTGAACAATGGAACAACCTTCAGACACAGAAATAACACCACTTCGTTTGTAGTA 1432
QY 661 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 720
Db 1433 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 1492
QY 721 AGCAATCTATAGACACACACAGCTCCCAATGATGAGTGGGTAAACCTCCACACGA 780
Db 1493 AGCAATCTATACACACACACCTCCCAATGATGAGTGGGTAACTCTCCACACAA 1552
QY 781 ATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTACTCTCAGGCTATCATTTGTTG 840
Db 1553 ATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTACTCTCAGGCTATCATTTGTTG 1612
QY 841 AATGGCTCTTCAGAACTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 1613 AATGGCTCTTCAGAACTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1672
QY 901 ACTGAACAAGATTTATCAATCATCTGCTAGCTTAAGCCCCACACAAAGAGTACCCATT 960
Db 1673 ACTGAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAACAAAGAGTACCCATT 1732
QY 961 CTTCCTTTTGTATCAGACGAGGAGTGTAGGACATAGTGTAGTGTAGTGTAGTGTAGTGT 1020
Db 1733 CTTCCTTTTGTATCAGACGAGGAGTGTAGGACATAGTGTAGTGTAGTGTAGTGTAGTGT 1782
QY 1021 ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 1793 ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1852
QY 1081 GTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAATCTCCCTAGCAGAGTAGTCTCT 1140
Db 1853 GTCCGCACTCCCTGGTCACTTGGCAAGATCAACTTAATCTCCCTAGCAGAGTAGTCTCT 1912
QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGGAACTCTGTTTATTTTA 1200

Db 1913 CAAAATCGAAGAGCTTTAGACTTGTAAACCGTGAAGAGGGGAACCTGTTTATTTTA 1972
QY 1201 GGAGGAAGAACGCTGTTTATTTATGTAATCAATCAGAAATGTCTCACTGAGAAAGTTAAAGAA 1260
Db 1973 GGGGAAGATGCTGTTTATTTATGTAATCAATCGGNAATCGTCACTGAGAAAGTTAAAGAA 2032
QY 1261 ATTGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAAACACCGAACGCTGGGCGCTC 1320
Db 2033 ATTGAGATCGAATACAACTAGAGCAGAGAGCTTCAAAAACACCGAACGCTGGGCGCTC 2092
QY 1321 CTCAGCAATGATGCCCTGGCTTCTCCCTCTTCTAGGACCTCTAGCAGCTCTAAATATTG 1380
Db 2093 CTCAGCAATGATGCCCTGGATTTCTCCCTCTTCTAGGACCTCTAGCAGCTCTAAATATTG 2152
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTTAAAGTTTCTCTTCCAGAATT 1440
Db 2153 CTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTTAACTCTCTTCTCTCCAGAATC 2212
QY 1441 GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAAACCCCA 1481
Db 2213 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCA 2253
RESULT 3
CR605851 2748 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DE012VJ24 of Placenta of Homo sapiens
DEFINITION (human).
ACCESSION CR605851
VERSION CR605851.1 GI:50486658
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2748)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2748)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1. .2748
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012VJ24"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 90.1%; Score 1333.8; DB 3; Length 2748;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 1 ATGCGCTCTCTTATCATCTTTTCTCTTACTTGTCTTACCCCTTTCGCTCTCACT 60
Db 772 ATGCGCTCTCTTATCATATTTTCTCTTACTGTCTTTTACCTCTTTCACCTCACT 831
QY 61 GCACCCCTCCATGCTGCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120

```
Db      ||||| 832 GCACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCAAGAGTTCTTATGGAGA 891
Qy      ||||| 121 ACGCGCTTCTGGAATATTGATGCCCATCATATAGAGTTTATTAAGGAAACTCC 180
Db      ||||| 892 ATGCAAGGTCGCGGAATATTGATGCCCATCGTATAGAGTCTTCTTAAGGAACCCCC 951
Qy      ||||| 181 ACCTTCACTGCCACACCCATATGCGCGCACTGCTATACTGCCACCTTTGCAATG 240
Db      ||||| 952 ACCTTCACTGCCACACCCATATGCGCGCACTGCTATACTGCCACCTTTGCAATG 1011
Qy      ||||| 241 CATGCAAACTACTCATTTATGGAACAGGAAATGATTAATCTTAGTTGCTGGAGACTT 300
Db      ||||| 1012 CATGCAAACTACTCATTTATGGAACAGGAAATGATTAATCTTAGTTGCTGGAGACTT 1071
Qy      ||||| 301 GGAGCCACTGTCTGTGGACTTACTTCAACCCATACAGATATGCTGTGAGTGGGGTGAATTT 360
Db      ||||| 1072 GGAGTCACTGTCTGTGGACTTACTTCAACCCATACAGATATGCTGTGAGTGGGGTGAATTT 1131
Qy      ||||| 361 CAAGTCAAGCAAGAAACAACTTAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
Db      ||||| 1132 CAAGTCAAGCAAGAAACAACTTAAGGAAGCAATCTCCCACTGACCCGGGTACTAT 1191
Qy      ||||| 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db      ||||| 1192 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC 1251
Qy      ||||| 481 CATACTCGCTCGTGAGCTTATTAATACACCTCACTCGGCTCATGAGTCTCAGCC 540
Db      ||||| 1252 CATACTCGCTCGTGAGCTTATTAATACACCTCACTCGGCTCATGAGTCTCAGCC 1311
Qy      ||||| 541 CAAACCCCTACTAATCTGTGGATGCGCTCCCGCTGCACCTCAGGCCATACATTTCAATC 600
Db      ||||| 1312 CAAACCCCTACTAATCTGTGGATGCGCTCCCGCTGCACCTCAGGCCATATGTTTCAATC 1371
Qy      ||||| 601 CTGTTCTCTGAACTAAGCAACTTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
Db      ||||| 1372 CTGTTACTCTGAACTAAGCAACTTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 1431
Qy      ||||| 661 GGACCTCTGTTTCCAACTCGGAATAAACCATACCTCAAACTGACCTGCTGTGTAATTT 720
Db      ||||| 1432 GGACCTCTGTTTCCAACTCGGAATAAACCATACCTCAAACTGACCTGCTGTGTAATTT 1491
Qy      ||||| 721 AGCAATACTATAGACAACACCACTCCCAATGATCAGGTGGTAAACCTCCACACGA 780
Db      ||||| 1492 AGCAATACTATAGACAACCACTCCCAATGATCAGGTGGTAAACCTCCACACGA 1551
Qy      ||||| 781 ATAGTCTGCTACCTCAGGAATATTTTGTGTGTGTTACCTCAGCCTATCATTTGTTG 840
Db      ||||| 1552 ATAGTCTGCTACCTCAGGAATATTTTGTGTGTGTTACCTCAGCCTATCATTTGTTG 1611
Qy      ||||| 841 AATGGCTCTCAGAACTATGTGCTTCCCTCATCTTCTAGTGGCCCTATGACCATCTAC 900
Db      ||||| 1612 AATGGCTCTCAGAACTATGTGCTTCCCTCATCTTCTAGTGGCCCTATGACCATCTAC 1671
Qy      ||||| 901 ACTGAACAGATTTATACAACTATGTCGTACCTTAAGCCCAACAACAAAGAGTACCATT 960
Db      ||||| 1672 ACTGAACAGATTTATACAGTTATGTCATCTTAAGCCCGCAACAAAGAGTACCATT 1731
Qy      ||||| 961 CTTTCTTTTGTATCAGACGAGGTGCTAGGCAGACTAGGTACTGGCATTTGGCGGTATC 1020
Db      ||||| 1732 CTTTCTTTTGTATAGAGCAGGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATC 1791
Qy      ||||| 1021 ACAACCTCTACTCAGTTACTACAACTATCTCAAGAAATAAATGTCACATGGAACAG 1080
Db      ||||| 1792 ACAACCTCTACTCAGTTACTACAACTATCTCAAGAAATAAATGTCACATGGAACAG 1851
Qy      ||||| 1081 GTCACCTGACTCCCTGGTCACTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Db      ||||| 1852 GTGGCGACTCCCTGGTCACTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT 1911
Qy      ||||| 1141 CAAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGGAACCTGTTTATTTTAA 1200
```

```
Db      1912 CAAAATCGAAGAGCTTTAGACTTTGCTAAACGCTGAAAGAGGGGGAACCTGTTTATTTTAA 1971
Qy      1201 GGAGAAGAACGCTGTTATTTATTTATGTAATCAATCCAGAAATTCCTACTGAGAAAGTTAAAGAA 1260
Db      1972 GGGAGAAGATGCTGTTATTTATTTATGTAATCAATCCGGAATTCGCTACTGAGAAAGTTAAAGAA 2031
Qy      1261 ATTTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCGCAACCTGGGGGCTC 1320
Db      2032 ATTTCGAGATCGAATACAACTAGAGCAGAGAGCTTCAAAACACCTGGACCCCTGGGGGCTC 2091
Qy      1321 CTAGCCCAATGATGCCCTGGGTTCTCCCTCTCTTAGGACCTCTAGGAGCTCTTAATATTTG 1380
Db      2092 CTCAGCCCAATGATGCCCTGGGTTCTCCCTCTCTTAGGACCTCTAGGAGCTCTTAATATTTG 2151
Qy      1381 TTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAAT 1440
Db      2152 CTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTGTAACTTTGTTCTCTTCCAGAAATC 2211
Qy      1441 GAAGCTCTAAAGCTTACAGATGGTCTTACAAATGGAACCCCA 1481
Db      2212 GAAGCTCTAAAGCTTACAAATGGAACCCCAAGATGAGTCCAA 2252

RESULT 4
CR625046 2748 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODI044YK06 of Placentia Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR625046
VERSION CR625046.1 GI:50505853
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (Bases 1 to 2748)
AUTHORS Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
        http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
        Faraday Avenue
        2 (Bases 1 to 2748)
REFERENCE 2. (Bases 1 to 2748)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
        BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
        - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
        end enriched, double-strand cDNA was digested with Not I and cloned
        into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
        was normalized. Library was constructed by Life Technologies, a
        division of Invitrogen.
FEATURES             Location/Qualifiers
     source           1..2748
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSODI044YK06"
                     /tissue_type="Placentia Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"

ORIGIN
Query Match      90.1%; Score 1333.8; DB 3; Length 2748;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy      1 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTTCTTACCCCTTTGCTCTCACT 60
Db      783 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTTCTTACCCCTTTTCACTCTCACT 842
Qy      61 GCACCCCTCCATGCTGCTGTACACCACTAGTCCCTTACCAAGAGTTTCTTATGAAGA 120
```


Qy	121	ACGGCTTCTCGAAATATTTGATGCCCCATCATATAGGAGTTTATCTAAGGGAAATCTCC	180
Db	897	ATGCAGCGTCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC	956
Qy	181	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG	240
Db	957	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG	1016
Qy	241	CATGCAAACTACTCATTTTGGACAGGAAATGATTAATCCTAGTTGTCTCTGGAGGACTT	300
Db	1017	CATGCAAACTACTCATTTTGGACAGGAAATGATTAATCCTAGTTGTCTCTGGAGGACTT	1076
Qy	301	GGAGCCACTGTCTGTTGGACTTACTTCACCCATACCAAGTAGTGTGATGGGGTGGAAATT	360
Db	1077	GGAGTCACTGTCTGTTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGTGGAGTT	1136
Qy	361	CAGGTCAGCGCAGAGAAAACAAAGTAAGAGNAGCAATCTCCCACTGACCCCGGGACAT	420
Db	1137	CAAGTCAGGCGAGAGAAAACATGTAAAAGAGTAATCTCCCAACTCACCCGGGTACAT	1196
Qy	421	AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC	480
Db	1197	GGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC	1256
Qy	481	CATACTCGCTGTGGAGCCTATTTAATACACCCCTCACTCGGCTCATGAGGTCTAGCC	540
Db	1257	CATACTCGCTGTGAAGCCTATTTAATACACCCCTCACTCGGCTCATGAGGTCTCGGCC	1316
Qy	541	CAAAACCTTACTTAACGTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC	600
Db	1317	CAAAACCTTACTTAACGTGTGGATGTGCTCCCTCGAACTTCAGGCCATATGTTTCAATC	1376
Qy	601	CCTGTTCTTGAACAATGGAAACAATTCAGACAGAAATAAACACCCTCTCCGTTTTAGTA	660
Db	1377	CCTGTACTTGAACAATGGAAACAATTCAGACAGAAATAAACACCCTCTCCGTTTTAGTA	1436
Qy	661	GGACCTCTTGTTTCCAACTCTGGAATAAACCATACCTCAAAACCTCCTGTGTAAATTTT	720
Db	1437	GGACCTCTTGTTTCCAACTCTGGAATAAACCCATACCTCAAAACCTCCTGTGTAAATTTT	1496
Qy	721	AGCAATACTATGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCACACGA	780
Db	1497	AGCAATACTATGACACAAACAGCTCCCAATGCATCAGGTGGGTAACTCTCCACACAA	1556
Qy	781	ATAGTCGCTTACCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG	840
Db	1557	ATAGTCGCTTACCCCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTTG	1616
Qy	841	AATGGCTCTTCAGAACTATGTGCTTCTCCTCATCTTAGTGCCCTATGACCATCTAC	900
Db	1617	AATGGCTCTTCAGAACTATGTGCTTCTCCTCATCTTAGTGCCCTATGACCATCTAC	1676
Qy	901	ACTGAAACAAGATTTATACAATCATGTCTGTAACCTTAAGCCCCACAACAAAGAGTACCCATT	960
Db	1677	ACTGAAACAAGATTTATACAGTTAIGTCATATCTAAGCCCCGCAACAAAGAGTACCCATT	1736
Qy	961	CTTCTTTTGTATCAGAGCAGGAGTGTCTAGGCAGACTAGGTACTGTGCATGCGAGTATC	1020
Db	1737	CTTCTTTTGTATCAGAGCAGGAGTGTCTAGGTGTCACTAGGTACTGTGCATGCGGTATC	1796
Qy	1021	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAATGCTGACATGGAAACAG	1080
Db	1797	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAAACGG	1856
Qy	1081	GTCACTGACTCCCTGGTCACTTGTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Db	1857	GTGCGCGACTCCCTGGTCACTTGTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1916
Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGGAACCTGTTATTTTATA	1200
Db	1917	CAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGGAACCTGTTATTTTATA	1976

Qy	1201	GGAGAAGAA	CGCTGTGTTATTAATGTTAATCAATCAGAAATGTTGCTACTGAGAAGATTAAAGAA	1266				
Db	1977	GGGGAAGAAT	GCCTGTTATTAATGTTAATCAATCCGGAAATCGTCTACTGAGAAGATTAAAGAA	2036				
Qy	1261	ATTCGAGATCGAAT	TACAACTGTAGACGAGAGGAGCTTCMAAACACCGAACGCTGGGGCCTC	1320				
Db	2037	ATTCGAGATCGAAT	TACAACTGTAGACGAGAGGAGCTTCGAACAACCTGGACCTCGGGCCTC	2096				
Qy	1321	CTCAGCCAATGGAT	TGCCCTGTCTCCCTTTTAGGACCTCTAGCAGCTCTTAATATTG	1380				
Db	2097	CTCAGCCAATGGAT	TGCCCTGTCTCCCTTTTAGGACCTCTAGCAGCTCTTAATATTG	2156				
Qy	1381	TTACTCTCTCTTGG	ACCTGTATCTTTTAACCTCCTGTTTAAGTTTCTCTCTCCAGAAATT	1440				
Db	2157	CTACTCTCTTGG	ACCTGTATCTTTTAACCTCCTGTTTAACCTTTGTCTCTCCAGAAATC	2216				
Qy	1441	GAAGCTGTAAAGCT	TACAGATGGTCTTACAAATGGAAACCCCA	1481				
Db	2217	GAAGCTGTAAACT	TACAAATGGAGCCCAAGATCAGTCCAA	2257				
RESULT 6								
BX391741/c								
LOCUS								
DEFINITION	BX391741	Homo sapiens	PLACENTA COT 25-NORMALIZED Homo sapiens cDNA	EST 29-APR-2004				
	clone CSODI051YM13	3-PRIME, mRNA	sequence.					
ACCESSION	BX391741							
VERSION	BX391741.1	GI:30619529						
KEYWORDS	EST.							
SOURCE	Homo sapiens							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.							
TITLE	Full-length cDNA libraries and normalization							
JOURNAL	Unpublished (2001)							
COMMENT	Contact: Genoscope							
Genoscope - Centre National de Sequencage								
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE								
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr								
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime								
end enriched, double-strand cDNA was digested with Not I and cloned								
into the Not I and EcoR V sites of the PCWSPORT 6 vector. Library								
was normalized. Library was constructed by Life Technologies, a								
division of Invitrogen. This sequence belongs to sequence cluster								
4215.r								
For more information about this cluster, see								
http://www.genoscope.cns.fr/cdna?S=CS0BA1022ZA10_CS02089_1&c=4215.r								
FEATURES								
source	Location/Qualifiers							
	1..935							
	/organism="Homo sapiens"							
	/mol_type="mRNA"							
	/db_xref="taxon:9606"							
	/clone="CS0DI051YM13"							
	/tissue_type="PLACENTA COT 25-NORMALIZED"							
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"							
	/notes="1st strand cDNA was primed with a NotI-oligo(dT)							
	primer. Five prime end enriched, double-strand cDNA was							
	digested with Not I and cloned into the Not I and EcoR V							
	sites of the PCWSPORT 6 vector. Library was normalized."							
ORIGIN								
Query Match	56.7%; Score 839; DB 5; Length 935;							
Best Local Similarity	94.2%; Pred. No. 5; 8e-237;							
Matches	880; Conservative	0; Mismatches	53; Indels	1; Gaps				
Qy	225	TGCCACTCTTTGATGATGCAAAATCTACTATTTGGACAGGAGAAATGATTAATCTAG	284					
Db	933	TGCCCTCTTTGATGATGCAAAATCTCA-TATTGGACAGAAATGATTAATCTCTAG	875					
Ov	285	TTGTCTCTGGAGGACTTTGGAGCCACTCTCTTTGGACTTACTTCACCCATACCAAGTATGTC	344					

Db	874	TTGCTCTGAGGACTTGAAGTCACTCTGTGGACTTACTTCAACCAACTGGTATGTC	815
Qy	345	TGATGGGGTGGAAATTAAGGTGAGCAAGAAACAAAGTAAGGAACCAATCTCCCA	404
Db	814	TGATGGGGTGGAGTTCAAGATCAGGCAAGAAACAAAGTAAGTAATCTCCCA	755
Qy	405	ACTGACCCGGGACATAGACACCCCTGACCCCTCAAGGACTAGTTCTCTCAAACTACA	464
Db	754	ACTACCCGGGTACATGGACCTCTAGCCCTCAAGGACTAGATCTCTCAAACTACA	695
Qy	465	TGAACCCCTCCGACCATATCTCCCTGTGGTGAACCTATTATTAACACCCCTACTCGGT	524
Db	694	TGAACCCCTCCGACCATATCTCCCTGTGGTGAACCTATTATTAACACCCCTACTCGGT	635
Qy	525	CGATGAGGTCTAGCCCAACCCCTACTACTGTGGATGTGCTCCCTGCATTCAG	584
Db	634	CGATGAGGTCTAGCCCAACCCCTACTACTGTGGATGTGCTCCCTGCATTCAG	575
Qy	585	GCCATACATTTCAATCCCTGCTTCTGAACAATGGAACAACCTTCAGACAGAAATAACAC	644
Db	574	GCCATATGTTTCAATCCCTGCTTCTGAACAATGGAACAACCTTCAGACAGAAATAACAC	515
Qy	645	CACCTCCGTTTATAGGACCTCTGTGTTCCCAATCTGGAATAACCCATACCTCAAACT	704
Db	514	CACCTCCGTTTATAGGACCTCTGTGTTCCCAATCTGGAATAACCCATACCTCAAACT	455
Qy	705	CACCTGTGTAATTTAGCAATATATAGACACACACAGCTCCCAATGCATCAGTGGGT	764
Db	454	CACCTGTGTAATTTAGCAATATATAGACACACACAGCTCCCAATGCATCAGTGGGT	395
Qy	765	AACACCTCCACAGCATGCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC	824
Db	394	AACCTCCACACAAATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC	335
Qy	825	AGCTATCATTTGTAATGGCTCTTCAGAACTATGTGCTCTCTCTCATTTCTTAGTGC	884
Db	334	AGCTATCATTTGTAATGGCTCTTCAGAACTATGTGCTCTCTCTCATTTCTTAGTGC	275
Qy	885	CCCTATGACCATCTACAGCAAGATTTATACATCATGTGCTTACCTTAAGCCCAAA	944
Db	274	CCCTATGACCATCTACAGCAAGATTTATACAGTTATGTATCTAAGCCCAAA	215
Qy	945	CAAAAGTACCCATCTCTCTTTGTTATCAGACAGGAGTGTAGGACAGACTAGTAC	1004
Db	214	CAAAAGTACCCATCTCTCTTTGTTATCAGACAGGAGTGTAGGACAGACTAGTAC	155
Qy	1005	TGGCATTTGGGAGTATCAACACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAA	1064
Db	154	TGGCATTTGGGAGTATCAACACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAA	95
Qy	1065	TGTGACATGGAACAGGTCTACTGCTCCCTGGTCACTTGCAGAGTCACTTACCTCCT	1124
Db	94	TGGGACATGGAACAGGTCTGCTGCTCCCTGGTCACTTGCAGAGTCACTTACCTCCT	35
Qy	1125	AGCAGCATGCTCTTCAAAATCGAAGCTTTA	1158
Db	34	AGCAGCATGCTCTTCAAAATCGAAGCTTTA	1

RESULT 7
 BX347952
 LOCUS
 DEFINITION BX347952 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 clone CSODI051Y13 5-PRIME, mRNA sequence.
 ACCESSION BX347952
 VERSION BX347952.1 GI:30375235
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 955)

AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?e=CS0BAE0232F07_AB02167_1&c=4215.r
FEATURES	Location/Qualifiers 1..955 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODI051Y13" /tissue_type="PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	Query Match 54.9%; Score 813.6; DB 5; Length 955; Best Local Similarity 93.7%; Pred. No. 2e-229; Matches 890; Conservative 0; Mismatches 56; Indels 4; Gaps 4; 226 GCCACTCTTTGATGATGCAATCACTTATTTGACAGGAAATGATTATCTTAGT 285 1 GCCACTCTTTGATGATGCAATCACTTATTTGACAGGAAATGATTATCTTAGT 60 286 TGTCTCTGGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCATACAGTATGTCT 345 61 TGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCATACAGTATGTCT 120 346 GATGGGGTGGAAATCAAGTCAAGCAAGAGAAAAACAAGTAAGAAACAATCTCCCA 405 121 GATGGGGTGGAGTTCAAGATCAGGCAAGAAAAACAATGTAAGAAAGTAATCTCCCA 180 406 CTGACCCGGGACATAGCACCCTAGCCCTACAGAGGACTAGTCTCTCAAACTACAT 465 181 CTCACCCGGGTACATGGACCTCTAGCCCTACAGAGGACTAGTCTCTCAAACTACAT 240 466 GAAACCTCTCGTACCATCTCTCGCTGGTGGCTATTTAATACCACTCCTCCTCGGCTC 525 241 GAAACCTCTCGTACCATCTCTCGCTGGTGGCTATTTAATACCACTCCTCCTCGGCTC 300 526 CATGAGGTCTCAGCCCAAAACCTTAACTGTTGATGTGCTCCCTCTGCACTTCAGG 585 301 CATGAGGTCTCAGCCCAAAACCTTAACTGTTGATGTGCTCCCTCTGCACTTCAGG 360 586 CCATACATTTCAATCTCTGTTCTGACATGCAACACTTTCAGCAGAGAAATAAACACC 645 361 CCATATGTTTCAATCTCTGTTCTGACATGCAACACTTTCAGCAGAGAAATAAACACC 420 646 ACTTCTCGTTTATAGTACGCTCTGTTTTCCTCAATCTGGAATAACCACTACCTCAAACTC 705 421 ACTTCTCGTTTATAGTACGCTCTGTTTTCCTCAATCTGGAATAACCACTACCTCAAACTC 480 706 ACCTGTGTAAATTTAGCAATATATAGACACAAACAGCTCCCAATGATCAGTGGGTA 765 481 ACCTGTGTAAATTTAGCAATATATAGACACAAACAGCTCCCAATGATCAGTGGGTA 540 766 ACACCTCCACAGCAATAGTCTGCTACCTCCTCAGGATATTTTGTCTGCTGCTACCTCA 825 541 ACTCCT-CCAAACAAATAGTCTGCTACCTCCTCAGGATATTTTGTCTGCTGCTGCTACCTCA 599

QY 826 GCCTATCATTTGTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATTTTAGTGCCC 885
Db 600 GCCTATCGTTGTTGAATGGCTCTTCAGAACTATGTGCTTCCTCTCATTTTAGTGCCC 659
QY 886 CCTATGACCATCTACACTGAACAAGATTTATACATCATGTGCTACCTTAAGCCCAAC 945
Db 660 CCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTATCATATCTTAAGCCCGCAAC 719
QY 946 AAAAGAGTACCATCTTCCTTTGTTATCAGAGCAGGAGTGCTAGGCAGACTAGGTACT 1005
Db 720 AAAAGAGTACCATCTTCCTTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACT 779
QY 1006 GGCATTGGCAGTATCACAACTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAT 1065
Db 780 GGCATTGGCGGTATCACAACTCTACTCAGTTCTACTACANACTATCTCAAGAACTAAT 839
QY 1066 GGTGACAT-GGAACAGGTCACTGACTCCTCGG-TCACCTTGCAGATCAACTTAACCTCC 1123
Db 840 GGGGACATGGGAACGGTTCCGCGACTTCCTGGTTACCTTGCAGATCAACTTAACCTCC 899
QY 1124 TAGCAGCAGTAG-TCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGC 1172
Db 900 TAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGC 949

RESULT 8
BX389656 931 bp mRNA linear EST 29-APR-2004
LOCUS BX389656 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1051YM13 5-PRIME, mRNA sequence.
ACCESSION BX389656
VERSION BX389656.1 GI:30462930
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 931)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAP0272B05_AF02531_1&c=4215.r

FEATURES
source Location/Qualifiers
1. .931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 50.8%; Score 752.2; DB 5; Length 931;
Best Local Similarity 92.3%; Pred. No. 3.3e-211;
Matches 856; Conservative 0; Mismatches 65; Indels 6; Gaps 6;

QY 267 GAAATGATTAACTCCTAGTTGCTCGAGGACTTTGGAGCCACTGTCTGTTGGACTTACTT 326
Db 1 GAAAAAGATTAACTCCTAGTTGCTCGAGGACTTTGGAGTCACTGTCTGTTGGACTTACTT 60
QY 327 CACCCCATACCCAGTATGCTGATGGGGGTGGAATTCAGGTTCAGGCAAGAGAAAAACAAGT 386
Db 61 CACCCAAACTGTTGATGCTGATGGGGGTGAGTTCAAGATCAGGCAAGAAAAACAATGT 120
QY 387 AAAGAGAGCAATCTCCCACTGACCCGGGGGACATAGCACCCCTAGCCCTTACAAAGGACT 446
Db 121 AAAGAGAGTAACTCCCACTGACCCGGGGTACATGTCACCTCTAGCCCTTACAAAGGACT 180
QY 447 AGTTCTCTCAAACTACATGAAACCTCCGTACCCATATCTCGCTCTGCTGAGCCCTATTAA 506
Db 181 AGATCTCTCAAACTACATGAAACCTCCGTACCCATATCTCGCTCTGCTGAGCCCTATTAA 240
QY 507 TACCACCTCCTCAGTCCGCTCCATGAGGTCTCAGCCCAAAACCTCTACTAACTGTTGGATG 566
Db 241 TACCACCTCCTCAGTCCGCTCCATGAGGTCTCGGCCCAAAACCTCTACTAACTGTTGGATG 300
QY 567 CTCCCTCCTCAGTCTCAGGCAATATTTCAATCCCTGTTCTGAACTGGAACAACTT 626
Db 301 CTCCCTCCTGAACTCAGGCAATATGTTTCAATCCCTGTTCTGAACTGGAACAACTT 360
QY 627 CAGCAGAAATAAACACCACTTCCGTTTGTAGTAGGACCTCTTGTTCCTGAACTGGAACAACTT 686
Db 361 CAGCAGAAATAAACACCACTTCCGTTTGTAGTAGGACCTCTTGTTCCTGAACTGGAACAACTT 420
QY 687 AACCATACCTCAAACTCCTGTTGTAATTTAGCAATACTATAGACACCACTGCT 746
Db 421 AACCATACCTCAAACTCCTGTTGTAATTTAGCAATACTATAGACACCACTGCT 480
QY 747 CCAATGATCAGTGGGTGACCTCCACACGAATAGTCTGCTACCTCAGGAATATT 806
Db 481 CCAATGATCAGTGGGTGACCTCCACACGAATAGTCTGCTACCTCAGGAATATT 540
QY 807 TTTTGTCTGTGTTACCTCAGCTATCATGTTTGAATGGCTCTTTCAGAACTATGTCCTT 866
Db 541 TTTTGTCTGTGTTACCTCAGCTATCATGTTTGAATGGCTCTTTCAGAACTATGTCCTT 600
QY 867 CCTCTCATTTCTAGTGGCCCTATGACATCTACACTGAAACAGATTTATACATCATGT 926
Db 601 CCTCTCATTTCTAGTGGCCCTATGACCATCTACCTGAAACAGATTTATACATCATGT 660
QY 927 CGTACCTAAGCCCAACAAAGAGTACCACTTCTTCTTTTGTATCAGAGCAGGAGT 986
Db 661 CATATCTAAGCCCGCAACAAAGAGTACCACTTCTTCTTTTGTATCAGAGCAGGAGT 720
QY 987 GCTAGGAGACTAGGTACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTACTACAA 1046
Db 721 GCTAGGTGCACTAGGTACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTACTACAA 779
QY 1047 ACTATCTCAAGAAATAATGTTGAC-ATGGAACAGGTCACTGACT-CCCTGGTCACTTGT 1104
Db 780 ACTATCTCAAGAAATAATGTTGGAACAATGTAACGGGTGCGGCACTCCCTGGTCACTTTC 839
QY 1105 CAAGATCAACTTAAC-TCCCTTAGCAGCAGTA-GTCTTCTCAAAATCGAAG-AGCTTTAGAC 1161
Db 840 CGAGATCAACTTACTTCCCTTAGCAGCAGGAGTCTCTCAAAATCGAAGAGCTTTAGCC 899
QY 1162 TTGCTAACCCGCCAAAGAGGGGAACC 1188
Db 900 TTGCTAACCCGCTGCAGAGGGGGAAC 926

RESULT 9
BX326647/c 872 bp mRNA linear EST 07-APR-2004
LOCUS BX326647 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1051YM13 3-PRIME, mRNA sequence.
ACCESSION BX326647
VERSION BX326647.1 GI:30334578
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 872)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAD0102C11_AD00948_l&c=4215.r

FEATURES
source Location/Qualifiers
1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 50.5%; Score 747.2; DB 5; Length 872;
Best Local Similarity 92.3%; Pred. No. 9.7e-210;
Matches 806; Conservative 0; Mismatches 64; Indels 3; Gaps 2;
278 ATCTAGTGTCTGGAGGACTTGGAGCCACTGTCTTGGACTTACTTACCCATACCA 337
872 ATCTAGTGTCTGGAGGACTTGGAGCCACTGTCTTGGACTTACTTACCCATACCA 813
338 GTATGCTGATGGGGTGGATTCAAGTTCAGGACGAGAAACACAGTAAAGAACAA 397
812 GGTATGCTATGGGGTGGAGTTCAAGATCAGGACGAGAAACACAGTAAAGTAA 754
398 TCTCCCAACTGACCCGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAA 457
753 TCTCCCAACTGACCCGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAN 694
458 AACTACATGAACCCCTCCGTACCCACTACTCGCTGGTGAGCTATTAAATACCACTCA 517
693 AACTACATGAACCCCTCCGTACCCACTACTCGCTGGTGAGCTATTAAATACCACTCA 634
518 CTCGGCTCCATGAGGTCTACGCCAAACCTACTAAGTGTGAGTGGCTCCCTGTCG 577
633 CTGGGCTCCATGAGGTCTCGGCCCAACCCCTACTAAGTGTGAGTGGCTCCCTGTCG 574
578 ACTTCAGGCCATACATTTCAATCCTGTCTCTGAAACATGAAACAACTTCAGACAGAAA 637
573 ACTTCAGGCCATATGTTTCAATCCTGTACTCTGAAACATGAAACAACTTCAGACAGAAA 514
638 TAAACACCACTTCGGTTTATAGGACCTCTTGTTCCTCAATCTCGAAATACCCATACCT 697
513 TAAACACCACTTCGGTTTATAGGACCTCTTGTTCCTCAATCTCGAAATACCCATACCT 454
698 CAACACCTCACTGTGTAATAATTTAGCAATATATATAGACACAAACAGCTCCCAATGATCA 757
453 CAACACCTCACTGTGTAATAATTTAGCAATATATATAGCAATATATATAGCAATGATCA 394
758 GGTGGTAAACCTCCACAGCAATAGTCTGCTACCTACCTCAGGAATATTTTCTGCTGTG 817

Db 393 GGTGGTAACTCCTCCACACAAATAGTCTGCTACCTACCCCTCAGGAATATTTTGTCTGTG 334
Qy 818 GTACCTCAGCCTATCATTTTGAATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCT 877
Db 333 GTACCTCAGCCTATCGTTGTTGAATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCT 274
Qy 878 TAGTGGCCCTATGACCATCTACATGAAACAGATTTTATACATCATGTGCTGCTACCAAGC 937
Db 273 TAGTGGCCCTATGACCATCTACATGAAACAGATTTTATACATCATGTGCTGCTACCAAGC 214
Qy 938 CCCACAAAAGAGTACCATCTTCTCTTTGTTATCAGACGAGGAGTGTAGGACAGAC 997
Db 213 CCGGCAACAAAAGAGTACCATCTTCTCTTTGTTATAGAGGAGGAGTGTAGGAGTGCAC 154
Qy 998 TAGGTACTGGCATTGGCAGTATCACAACTCTACTCAGTTCTACTACTACAACTATCTCAAG 1057
Db 153 TAGGTACTGGCATTGGCAGTATCACAACTCTACTCAGTTCTACTACTACAACTATCTCAAG 94
Qy 1058 AATAAATGGTGACATGGAAACAGGTCATGACTCCCTGCTGCTGCTGCAAGATCAACTTA 1117
Db 93 AACTAAATGGGACATGGAAACAGGTCGCGGACTCC--TGGTCACCTGCAAGATCAACTTA 36
Qy 1118 ACTCCCTAGCAGCAGTAGTCTTCAAAATCGAGA 1152
Db 35 ACTCCCTAGCAGCAGTAGTCTTCAAAATCGAGA 1

RESULT 10
BX408752 903 bp mRNA linear EST 01-MAY-2004
BX408752 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120
5-PRIME, mRNA sequence.
ACCESSION BX408752
VERSION BX408752.2 GI:46931100
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 903)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30648825.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAA008ZH03_CS00697_2&c=4215.r

FEATURES
source Location/Qualifiers
1..903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013Y120"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
Query Match 50.0%; Score 741; DB 5; Length 903;

```
Best Local Similarity 93.0%; Pred. No. 6.8e-208;
Matches 796; Conservative 0; Mismatches 58; Indels 2; Gaps 2;

QY 4 GCCCTCCCTTATCATACACTTTCTTTTACTGTGTTCTTACCCCTTTTCGCTCTCACTGCA 63
Db 50 GGCCTCCCTTANATATTTTCTTTTACTGTGTTTACCTCTTTTCACTCTCACTGCA 109
QY 64 CCCCTCCCTGCTGTGTACACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGACG 123
Db 110 CCCCTCCCTGCTGTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGATG 169
QY 124 CGCTTCTCTGGAATATTTGATCCCATCATATAGGAGTTTATCTAAGGAACTCCACC 183
Db 170 GAGCTCCGGAATATTTGATCCCATCATATAGGAGTTTCTAAGGAACTCCACC 229
QY 184 TTCCTGCCACACCCATATGCCCCGCAACTGCTATAAATCTGCACTCTTTGATGAT 243
Db 230 TTCCTGCCACACCCATATGCCCCGCAACTGCTATAAATCTGCACTCTTTGATGAT 289
QY 244 GCAATATCTATTTGGACAGGAAATGATTAATCTAGTTGTCCTGGAGGACTTGA 303
Db 290 GCAATATCTATTTGGACAGGAAATGATTAATCTAGTTGTCCTGGAGGACTTGA 349
QY 304 GCACCTGTCTGTGGACTTACTTCCCATACAGTATGTTCTGATGGGGTGGAAATCAA 363
Db 350 GTCACTGTCTGTGGACTTACTTCCCATACAGTATGTTCTGATGGGGTGGAGTTCAA 409
QY 364 GTTCAGGCAAGAGAAACAAAGTAAAGAGCAATCTCTCCAACTGACCCGGGACATAGC 423
Db 410 GATCAGGCAAGAGAAACAAAGTAAAGAGCAATCTCTCCAACTGACCCGGGACATAGC 469
QY 424 ACCCTAGCCCTACAAAGGACTGTCTCTCAAACTACATGAACCTCCGTAACCAT 483
Db 470 ACTCTAGCCCTACAAAGGACTGTCTCTCAAACTACATGAACCTCCGTAACCAT 529
QY 484 ACTCGCTGTGAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGTCTTCAGCCCAA 543
Db 530 ACTCGCTGTGAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGTCTTCAGCCCAA 589
QY 544 AACCTTACTAACTGTGGATGTCCTCCCTGCACTTCAGGCTACATTTCAATCCCT 603
Db 590 AACCTTACTAACTGTGGATGTCCTCCCTGCACTTCAGGCTACATTTCAATCCCT 649
QY 604 GTTCTGAACTAGGAACAACTTCAGCAGAGAAATAAACACCACTTCCTGTTAGTAGA 663
Db 650 GTACCTGAACTAGGAACAACTTCAGCAGAGAAATAAACACCACTTCCTGTTAGTAGA 709
QY 664 CCTCTGTTTCCAACTCGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 723
Db 710 CCTCTGTTTCCAACTCGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 769
QY 724 AATCTATAGACAAACAGCTCCCAATGCATCAGTGGGTAAACCTCCACAGCAATA 783
Db 770 AATCTATACATACAAACCACTCCCAATGCATCAGTGGGTAACTCCTCCACCAATA 829
QY 784 GTCTGCTACCTCAGGAATATTTTGTGCTGTGGTACCTCAGCCTCATCTGTTTGAAT 843
Db 830 GTCTGCTACCTCAGG-ATAATTTTGTGCTG-GGTACCTCAGCCTATTCGTTTGGAT 887
QY 844 GGCTCTTCAGAACTCTA 859
Db 888 GGCTCTTCAGAACTCTA 903
```

```
RESULT 11
BX408733
LOCUS
DEFINITION BX408733 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24
5-PRIME, mRNA sequence.
ACCESSION BX408733
VERSION BX408733.1 GI:30635957
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 883)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAA0082E02_CS00686_l&c=4215.r
.
FEATURES
source
Location/Qualifiers
1..883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 49.8%; Score 737.2; DB 5; Length 883;
Best Local Similarity 93.2%; Pred. No. 9e-207;
Matches 825; Conservative 0; Mismatches 54; Indels 6; Gaps 5;
QY 29 TTACTGTTCTTCTTACCCCTTTTCGCTCTCACTGACACCCCTCCATGCTGCTGTACAACCA 88
Db 1 TTACTGTTCTTCTTACCCCTTTTCGCTCTCACTGACACCCCTCCATGCGCTGTATGACCA 58
QY 89 GTAGCTCCCTTACCAAGAGTTTCTATGAAGAACCGCGCTTCCTGGAAATATTGATGCC 148
Db 59 GTAGCTCCCTTACCAAGAGTTTCTATGAAGAACCGAGCGTCCCGGAAATATTGATGCC 118
QY 149 CATCATATAGAGTTTATCTAAGGAACTCCACCTTCACCTGCCACACCCCATATGCC 208
Db 119 CATCGTATAGAGTCTTTTAAAGGAAACCCCTTCACCTGCCACACCCCATATGCC 178
QY 209 GCAACTGCTATACTTGCCTCTTTGTCATGTCATGCAATACCTATTATTGGACAGGA 268
Db 179 GCAACTGCTATCACTTGCCTCTTTGTCATGTCATGCAATACCTATTATTGGACAGGA 238
QY 269 AAATGATTAACTCTAGTTGCTCTCGAGACCTTGAGGCCACTGTCTGTGGACTTACTTCA 328
Db 239 AAATGATTAACTCTAGTTGCTCTCGAGACCTTGAGGTCACTGTCTGTGGACTTACTTCA 298
QY 329 CCCATACAGATGTTGATGGGGTGGAAATTCAGGTTCAGGCAAGAGAAACAAAGTAA 388
Db 299 CCCAACTGGTATGTTGATGGGGTGGAGTTCAAGATTCAGGCAAGAGAAACAAAGTAA 358
QY 389 AGGAAGCAATCTCCCACTGACCCGGGACATAGCACCCCTAGCCCTTACAAAGGACTAG 448
Db 359 AAGAAGTAACTCTCCCACTGACCCGGGACATAGGCACTCTAGCCCTTACAAAGGACTAG 418
QY 449 TTCTCTCAAACTACATGAACCCCTCCGTACCCATACCTGCGCTGTGAGCCTATTTAATA 508
Db 419 ATCTCTCAAACTACATGAACCCCTCCGTACCCATACCTGCGCTGTGAGCCTATTTAATA 478
QY 509 CCACCTCACTCGGCTCCATGAGTCTCAGGCCCAACCCCTACTAACTGTTGGATGCC 568
Db 509 CCACCTCACTCGGCTCCATGAGTCTCAGGCCCAACCCCTACTAACTGTTGGATGCC 568
```

Db 479 CCACCCTCACTGGGCTCCATGAGGTCTCGCGCCAAACCCCTACTAACTGTGTGATATGCC 538
 Qy 569 TCCCCCTGCACCTTCAGGCCATACATTTCAATCCCTGTTCTGTGACCAATGGAACAATCTCA 628
 Db 539 TCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGGAACAATGGAACAATCTCA 598
 Qy 629 GCACAGAAATAACACACACTTCGGTTTGTAGTAGGACCTCTTGTTCCTCAATCTGGAATAA 688
 Db 599 GCACAGAAATAACACACACTTCGGTTTGTAGTAGGACCTCTTGTTCCTCAATCTGGAATAA 658
 Qy 689 CCCATACCTCAAACTCCTCACTGTGTAAATTTAGCAATACTATAGACACACACAGCTCCC 748
 Db 659 CCCATACCTCAAACTCCTCACTGTGTAAATTTAGCAATACTATAGACACACACAGCTCCC 718
 Qy 749 AATGATCA-GGTGGGTAAACACTCCACACGAAATAGTCTGCTTACCTCAGCAATATT 807
 Db 719 AATGATCAAGGTGGGTAAACACTCCACACGAAATAGTCTGCTTACCTCAGCAATATT 777
 Qy 808 TTTGTCTGTGTACCTCAGCTTATCATTTGTTGAATGGCTC-TTCAGAAATCTATGTGCTT 866
 Db 778 TTTGTCTGTGTACCTCAGCTTATCATTTGTTGAATGGCTC-TTCAGAAATCTATGTGCTT 837
 Qy 867 CCTCTCATTTCTAGTGGCCCC-TATGACCAATCTACACTGAACAAG 910
 Db 838 CCTCTCATTTCTAGTGGCCCC-TATGACCAATCTACACTGAACAAG 882

RESULT 12
 BX347619/c 877 bp mRNA linear EST 23-APR-2004
 LOCUS BX347619 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CSOD1022XJ18 3-PRIME, mRNA sequence.

ACCESSION BX347619.1 GI:30379218
 VERSION 1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 877)
 AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 4215.r

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0BAD0092B06_AD00811_1&c=4215.r

FEATURES
 source Location/Qualifiers
 1. .877
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1022XJ18"
 /tissue type="PLACENTA COT 25-NORMALIZED"
 /clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="First strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 49.2%; Score 728.4; DB 5; Length 877;
 Best Local Similarity 93.2%; Pred. No. 3.7e-204;
 Matches 803; Conservative 0; Mismatches 55; Indels 4; Gaps 4;

Qy 503 TTAATACCAACCCCTCACTCGGCTCCATGAGGTCTC-AGCCCAAAACCCCTACTAACTGTG 561
 Db 877 TAAACCAACCCCTCACTCGGCTCCATGAGTNCCTCGGGCCAAACCCCTACTAACTGTGG 818
 Qy 562 ATGTGCTCCCTGCACCTTCAGGCCATACATTTCAATCCCTGTTCTCGAACAATGGAAC 621
 Db 817 ATATGCTCCNCCTGAACCTTCAGGCCATATGTTTCAATCCCTGTACCTGAACNAATGGAAC 758
 Qy 622 AACTTCAGCACAGA-AATAAACACCACTTCCTGTTTGTAGTAGGACCTCTTGTTCCTCAATCT 680
 Db 757 ACTTTCAGCACAGANATTAACACCACTNCCGTTTGTAGTAGGACCTCTTGTTCCTCAATCT 698
 Qy 681 GGAATAAACCCTACCTCAACCTCCTGTGTAAATTTAGCAATACTATAGACACACAAC 740
 Db 697 GGAATAAACCCTACCTCAACCTCCTGTGTGT-AAAATTTAGCAATACTATAGACACAAC 639
 Qy 741 CAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGAAATAGTCTGCTACCTACCTCAGG 800
 Db 638 CAACTCNAATGATCAGGTGGGTAACTCTCTCCACACAAATAGTCTGCTACCTACCTCAGG 579
 Qy 801 AATATTTTGTGCTGTGGTACCTCAGCCTATCATTTGTTGAATGGCTCTTTCAGAAATCTAT 860
 Db 578 AATATTTTGTGCTGTGGTACCTCAGCCTATCATTTGTTGAATGGCTCTTTCAGAAATCTAT 519
 Qy 861 GTGCTTCTCTCATTTCTAGTGGCCCCCTATGACCACTCTACCTGAAACAAGATTTATACAA 920
 Db 518 GTGCTTCTCTCATTTCTAGTGGCCCCCTATGACCACTCTACCTGAAACAAGATTTATACAG 459
 Qy 921 TCATGTGTAACCTTAAGCCCAACAAAGAGTACCCATCTTCTCTTGTGTTATCAGAGC 980
 Db 458 TTAGTCATATCTTAAGCCCAACAAAGAGTACCCATCTTCTCTTGTGTTATAGGAGC 399
 Qy 981 AGGAGTGCTAGGAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
 Db 398 AGGAGTGCTAGGAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 339
 Qy 1041 CTACAAACTATCTCAAGAAATAAATGGTGCATGGAACAGGTGCATGCTGCTGCTGCTGCTGCT 1100
 Db 338 CTACAAACTATCTCAAGAAATAAATGGGACATGGAACGGTGCCTGCTGCTGCTGCTGCTGCT 279
 Qy 1101 CTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGA 1160
 Db 278 CTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGA 219
 Qy 1161 CTTGCTAACCGCAAGAGAGGGGACCTGTTATTTTATAGGAGAAACGCTGCTTTATTA 1220
 Db 218 CTTGCTAACCGCTGAAGAGGGGACCTGTTATTTTATAGGAGAAACGCTGCTTTATTA 159
 Qy 1221 TGTTAATCAATCAGAAATTTGCTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATG 1280
 Db 158 TGTTAATCAATCGGAATCTGCTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAACG 99
 Qy 1281 TAGAGCAGAGAGCTTCAAAACACCGAACCGCTGGGGCTCTCTCAGCAATGGATGCCCTG 1340
 Db 98 TAGAGCAGAGAGCTTCAAAACACCTGGACCTGGGGCTCTCTCAGCAATGGATGCCCTG 40
 Qy 1341 GGTTCCTCCCTTCTTAGGACCT 1362
 Db 39 GATTCCTCCCTTCTTAGGACCT 18

RESULT 13
 BX388766/c 836 bp mRNA linear EST 29-APR-2004
 LOCUS BX388766 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CSOD104YK06 3-PRIME, mRNA sequence.
 ACCESSION BX388766
 VERSION BX388766.2 GI:46877917
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 836)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30462383.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AU012ZB04_U01082_1&c=4215.r.
Location/Qualifiers

FEATURES

source

1..836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI044YK06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 48.2%; Score 714; DB 5; Length 836;
Best Local Similarity 92.6%; Pred. No. 6.8e-200;
Matches 772; Conservative 0; Mismatches 60; Indels 2; Gaps 2;

QY 565 TGCCTCCCTCCGCTTACAGGCATACATTTCAATCCCTGTTCCCTGAACAATGGAAACAC 624
DB 835 TGCTTCCCTCCGCTTACAGGCATACATTTCAATCCCTGTTCCCTGAACAATGGAAACAC 776
QY 625 TTCCAGCAGACAAA-TAAACACCACTTCCGTTTTAGTAGGACCTCTGTTTCCAAATCTGGA 683
DB 775 TTCCAGCAGACAAA-TAAACACCACTTCCGTTTTAGTAGGACCTCTGTTTCCAAATCTGGA 716
QY 684 AATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATACATATAGACACACACAG 743
DB 715 AATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATACATATAGACACACACAA 656
QY 744 CTCCCAATGCATCAGTGGGTAAACACCTCCACACGAATAGTCTGCCTACCTCAGGAAT 803
DB 655 CTCCCAATGCATCAGTGGGTAAACACCTCCACACGAATAGTCTGCCTACCTCAGGAAT 596
QY 804 ATTTTG 863
DB 595 ATTTTG 536
QY 864 CTTCCTCTCATCTTAGTGTGCTCCCTATGACCATCTACACTGACAGATTTTACAAATCA 923
DB 535 CTTCCTCTCATCTTAGTGTGCTCCCTATGACCATCTACACTGACAGATTTTACAGTTA 476
QY 924 TGTGCTACCTAGAGCCCAACAAAGAGTAGTACCATCTCTCTTTGTTTATCAGAGCAGG 983
DB 475 TGTGCTACCTAGAGCCCAACAAAGAGTAGTACCATCTCTCTTTGTTTATGAGGAGG 416
QY 984 AGTGCTAGGACATAGGTACTGGCATTTGGCAGTATCAACAACCTCTACTCAGTTCTACTTA 1043
DB 415 AGTGCTAGGACATAGGTACTGGCATTTGGCAGTATCAACAACCTCTACTCAGTTCTACTTA 356
QY 1044 CAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTACTCCCTGGTCACTTT 1103
DB 355 CAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTACTCCCTGGTCACTTT 296
QY 1104 GCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTT 1163

DB 295 GCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTT 236
QY 1164 GCTAACCGCCGAAAAGAGGGGAACTGTTATTTTAGGAGAGAAACGCTGTTATTATGT 1223
DB 235 GCTAACCGCTGAAAAGAGGGGAACTGTTATTTTAGGAGAGAAATGCTGTTATTATGT 176
QY 1224 TAATCAATCCAGAAATTTGCTACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACACGTAG 1283
DB 175 TAATCAATCCGAAATGCTACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACACGTAG 116
QY 1284 AGCAGAGAGCTTCAAAACACCGAACCTGGGGCTCTCTAGCCAAATGAGTGCCTGGGT 1343
DB 115 AGCAGAGAGCTTCAAAACACCTGGACCTGGGGCTCTCTAGCCAAATGAGTGCCTGGAT 56
QY 1344 TCTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATTGTTACTCTCTTTGGACC 1397
DB 55 TCTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATTGTTACTCTCTTTGGGCC 3

RESULT 14

EX408734

LOCUS

DEFINITION BX408734 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012J24
5-PRIME, mRNA sequence.

ACCESSION

BX408734

VERSION

BX408734.1

GI:30635959

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Eukarya; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 921)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

was not normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0BAA008ZB02_CS00686_2&c=4215.r

FEATURES

source

Location/Qualifiers

1..921

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DE012J24"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

ORIGIN

Query Match 47.9%; Score 709.8; DB 5; Length 921;
Best Local Similarity 88.6%; Pred. No. 1.2e-198;
Matches 815; Conservative 0; Mismatches 98; Indels 7; Gaps 4;

QY 29 TTACTGTTCTCTTACCCCTTTTCCGCTCTCACTGCACCCCTCCATGCTGCTGTACACCA 88
DB 1 TTCTGTTCTTTTACCCCTCTTCTCACTGCTCACTGCACCCCTCCATGCTGCTGTACCA 60
QY 89 GTAGCTCCCTTTACCAAGAGTTTCTATGAGAAACGCGGCTTCTCTGGAAATATTGTATGCC 148


```
QY 1100 CTTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAG 1159
|||
Db 237 CTTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAG 178
|||
QY 1160 ACTTCTTAACCCCAAAAGAGGGGAACTCTGTTATTTTATTTAGGAGAAAGACGCTGTTATT 1219
|||
Db 177 ACTTCTTAACCCCTGAAGAGGGGAACTCTGTTATTTTATTTAGGAGAAAGATGCTGTTATT 118
|||
QY 1220 ATGTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAAT 1279
|||
Db 117 ATGTTAATCAATCCGAAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAAC 58
|||
QY 1280 GTAGAGCAGAGAGCTTCAAAACCGAAGCTGCGGCTCC 1321
|||
Db 57 GTAGAGCAGAGAGCTTCAAAACACTGGACCTGGGGGCGGCC 16
|||
```

```
RESULT 16
BX367907 828 bp mRNA linear EST 28-APR-2004
LOCUS BX367907 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI051YM13 5-PRIME, mRNA sequence.
```

```
ACCESSION BX367907
VERSION BX367907.1 GI:30459586
```

```
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
1 (bases 1 to 828)
```

```
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
```

```
Full-length cDNA libraries and normalization
```

```
Unpublished (2001)
```

```
Contact: Genoscope
```

```
Genoscope - Centre National de Sequencage
```

```
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
```

```
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
```

```
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
```

```
end enriched, double-strand cDNA was digested with Not I and cloned
```

```
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

```
was normalized. Library was constructed by Life Technologies, a
```

```
division of Invitrogen. This sequence belongs to sequence cluster
```

```
4215.r
```

```
For more information about this cluster, see
```

```
http://www.genoscope.cns.fr/cdna?s=CS0AT022ZG01_T02023_1&c=4215.r.
```

```
Location/Qualifiers
```

```
1. 828
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="CS0DI051YM13"
```

```
/tissue_type="PLACENTA COT 25-NORMALIZED"
```

```
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
```

```
/note="1st strand cDNA was primed with a NotI-oligo (dT)
```

```
primer. Five prime end enriched, double-strand cDNA was
```

```
digested with Not I and cloned into the Not I and EcoR V
```

```
sites of the pCMVSPORT 6 vector. Library was normalized."
```

```
ORIGIN
```

```
Query Match 47.5%; Score 703.2; DB 5; Length 828;
```

```
Best Local Similarity 93.8%; Pred. No. 1.1e-196;
```

```
Matches 754; Conservative 0; Mismatches 48; Indels 2; Gaps 2;
```

```
110 TTCTATGAAGACCGCGCTCTCGAAATATTGATGCCCCCATATAGAGCTTTATCTTA 169
```

```
9 TTTCTTGGAGATGCGAGCGTCCGG-ATAATTGATGCCCATCGTAGGAGCTTTCTTA 67
```

```
170 AGGGAACCTCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATACTCTGCCA 229
```

```
68 AGGGAACCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATACTCTGCCA 127
```

```
230 CTCTTTGCATGCATGCAAACTCATATTATGGACAGGGAATGATTAATCTTAGTTGTC 289
```

```
Db 128 CTCTTTGCATGCATGCAAACTCATATTATTGGACAGGAAAAATGATTAATCTTAGTTGTC 187
|||
QY 290 CTGGAGACTTTGGAGCCCACTGTCTGTTGGACTTACTTCCCAATACCAGATGATGCTGATG 349
|||
Db 188 CTGGAGACTTTGGAGTCACTGTCTGTTGGACTTACTTCCCAATACCAGATGATGCTGATG 247
|||
QY 350 GGGGTGGAATTCAGGTTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGA 409
|||
Db 248 GGGGTGAGTTCAGATCAGGCAAGAAAAACAATGATAAAGAGAGTAAATCTCCCAACTCA 307
|||
QY 410 CCGGGGACATAGACACCCCTAGCCCTTAAAGAGACTAGTTCTCTCAAAACTACATGAAA 469
|||
Db 308 CCGGGTACATAGCACCTCTAGCCCTTCAAAAGACTAGATCTCTCAAAACTACATGAAA 367
|||
QY 470 CCCTCCGTAGCCATACCTCGCTGAGCGCTATTTAATACCCCTCACTCGGCTCCATG 529
|||
Db 368 CCCTCCGTAGCCATACCTCGCTGAGCGCTATTTAATACCCCTCACTCGGCTCCATG 427
|||
QY 530 AGGTCTCAGCCCAAAACCTTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAGGCCAT 589
|||
Db 428 AGGTCTCGGCCCAAAACCTTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAGGCCAT 487
|||
QY 590 ACATTTCAATCCCTGTTCTCGAACATGGAACAACATTCAGCAGAGAAATAAACCACCTT 649
|||
Db 488 ATGTTTCAATCCCTGTTCTCGAACATGGAACAACATTCAGCAGAGAAATAAACCACCTT 547
|||
QY 650 CCGTTTTAGTAGGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACCT 709
|||
Db 548 CCGTTTTAGTAGGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACCT 607
|||
QY 710 GTGTAAAATTTAGCAATATATATAGACACAACACCTCCCAATGATCAGTGGGTAAACAC 769
|||
Db 608 GTGTAAAATTTAGCAATATATATAGACACAACACCTCCCAATGATCAGTGGGTAAACAC 667
|||
QY 770 CTCCACACGAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGTTACCTCAGCT 829
|||
Db 668 CTCCACACGAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGTTACCTCAGCT 727
|||
QY 830 ATCAATGTTTGAATGGCTCTTTCAGAACTATGCTCTCTCTCATTTCTAGTGCCCCCTA 889
|||
Db 728 ATCGTGTGTTGAATGGCTC-TCAGAACTATGCGCTCTCTCATTTCTTAGGGCCCCCTA 786
|||
QY 890 TGACCATCTACACTGAACAAGATT 913
|||
Db 787 TGACCATCTACACTGAACAAGATT 810
|||
```

```
RESULT 17
```

```
BX368078/c
```

```
LOCUS
```

```
DEFINITION
```

```
ACCESION
```

```
VERSION
```

```
KEYWORDS
```

```
SOURCE
```

```
ORGANISM
```

```
REFERENCE
```

```
AUTHORS
```

```
TITLE
```

```
JOURNAL
```

```
COMMENT
```

```
BX368078 846 bp mRNA linear EST 26-APR-2004
BX368078 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI051YM13 3-PRIME, mRNA sequence.
```

```
BX368078
```

```
BX368078.2 GI:46572908
```

```
EST.
```

```
Homo sapiens (human)
```

```
Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
1 (bases 1 to 846)
```

```
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
```

```
Full-length cDNA libraries and normalization
```

```
Unpublished (2001)
```

```
On May 8, 2003 this sequence version replaced gi:30447710.
```

```
Contact: Genoscope
```

```
Genoscope - Centre National de Sequencage
```

```
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
```

```
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
```

```
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
```

```
end enriched, double-strand cDNA was digested with Not I and cloned
```

```
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

```
was normalized. Library was constructed by Life Technologies, a
```

```
division of Invitrogen. This sequence belongs to sequence cluster
```

4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AU012D05_U01092_1&c=4215.r.

FEATURES

Location/Qualifiers
1. .846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 47.2%; Score 698.6; DB 5; Length 846;
Best Local Similarity 92.8%; Pred. No. 2.5e-195;
Matches 786; Conservative 0; Mismatches 56; Indels 5; Gaps 5;

QY 467 AAACCTCCGTACCCATACCTCGCTGGTGGAGCCTATTTTAATACCAC-CCTCACTCGGCTC 525
DB |||||
QY 846 AAACCTCCGTACCCATACCTCGCTGGTGGAGCCTATTTTAATACAAACCTCACTGGCTC 787
DB |||||

QY 526 CATGAGGTCTACGCC-AAACCCCTACTACTCTTGGATGTGCTCCCTCGACCTTCA- 583
DB |||||

QY 786 ATATAGGTCTCGGCCCAAAACCCCTACTACTGTGAATATGCTCCCTCGAATTCAG 727
DB |||||

QY 584 GGCCATACATTTCAAT-CCCTGTTCTCGAAATGGAACAACTTCAGCAGAGAAATAAAC 642
DB |||||

QY 726 GGCCATATGTTTCAATNCCCTGTACCTGANCATGNAACAACTTCAGCAGAGAAATAAAC 667
DB |||||

QY 643 ACCACTTCCGTTTGTAGTAGGACCTCTGTTTCCAAATCTGGAATTAACCCATCACTCAAC 702
DB |||||

QY 666 ACCACTNCCGTTTGTAGTAGGACCTCTGTTTCCAAATCTGGAATTAACCCATCACTCAAC 607
DB |||||

QY 703 CTACCTGTGTAATTTAGCAATATATAGACACACACAGCTCCCAATGCAATCACTAGTGG 762
DB |||||

QY 606 CTCACCTGTGTAATTTAGCAATATATAGACACACACAGCTCCCAATGCAATCACTAGTGG 547
DB |||||

QY 763 GTAACACTCCACAGCAATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGTACC 822
DB |||||

QY 546 GTAACCTCTCCACACAAATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGTACC 487
DB |||||

QY 823 TCAGCCTATCATTTGTTGAATGGCTCTTCAGAAATCTAATGTCTCTCTCATTTTGTAGT 882
DB |||||

QY 486 TCAGCCTATCGTTGTTGAATGGCTCTTCAGAAATCTAATGTCTCTCTCATTTTGTAGT 427
DB |||||

QY 883 CCCCCTATGACCATCTACCTGAACAGATTTATACATCATGTCTGTACCTAAGCCCCAC 942
DB |||||

QY 426 CCCCCTATGACCATCTACCTGAACAGATTTATACATCATGTCTGTACCTAAGCCCCCGC 367
DB |||||

QY 943 AACAAAGAGTACCCTATCTCTCTTTGTTATACAGACAGAGTGTAGGACAGACTAGGT 1002
DB |||||

QY 366 AACAAAGAGTACCCTATCTCTCTTTGTTATAGGACAGAGTGTAGGACAGACTAGGT 307
DB |||||

QY 1003 ACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAAATA 1062
DB |||||

QY 306 ACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAAATA 247
DB |||||

QY 1063 AATGGTGAATGGAACAGGTCACTGACTCCCTGTGTCACTTGGCAAGATCAACTTAATCTC 1122
DB |||||

QY 246 AATGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAATCTC 187
DB |||||

QY 1123 CTACGACGAGTGTCTTCAAAATCGAGAGCTTTAGACTTGTAAACCGCCAAAGAGGG 1182
DB |||||

QY 186 CTACGACGAGTGTCTTCAAAATCGAGAGCTTTAGACTTGTAAACCGCTGGAAGAGGG 127
DB |||||

QY 1183 GGAACCTGTTTATTTTGGAGAGAGAACCGCTGTTTATTTATTTATTAATCAATCCAGAAATGTC 1242
DB |||||

QY 126 GGAACCTGTTTATTTTGGAGAGAGATGCTGTTTATTTATTTATTAATCAATCCGGAATCGTC 67
DB |||||

QY 1243 ACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGACAGAGAGCTTCAAAAC 1302
DB |||||

QY 66 ACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAAGT-GAGCAGAGAGGCTTCAAAAC 8
DB |||||

QY 1303 ACCGAGAC 1309
DB 7 ATGGACC 1

RESULT 18

BX389657
LOCUS
DEFINITION
BX389657 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION
BX389657
VERSION
BX389657.2 GI:46875146
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 995)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On May 8, 2003 this sequence version replaced gi:30462931.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAF0272B05_AF02531_2&c=4215.r

FEATURES

Location/Qualifiers
1. .995
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 47.0%; Score 696.2; DB 5; Length 995;
Best Local Similarity 86.8%; Pred. No. 1.4e-194;
Matches 826; Conservative 0; Mismatches 114; Indels 12; Gaps 5;

QY 266 GGAATATGATTAATCTCTAGTTCCTGGAGAGCTTGGAGCCTGTCTGTGGACTTACT 325
DB |||||

QY 326 TCACCCATACAGTATCTCTGATGGGGTGGAAATTCAGGTTCAGGCAAGAGAAAAACAAG 385
DB |||||

QY 60 TCACCCAAATGTTATGTTCTGATGGGGTGGAGTTCAAGATCAGGCAAGAAAAACAATG 119
DB |||||

QY 386 TAAAGGAAGCAATCTCCCACTGACCCGGGAGACATAGACCCCTAGCCCTTACAAAGAC 445
DB |||||

QY 120 TAAAGGAAGTAACTCTCCCACTCACCCTGGGTACATGGCACCTCTAGCCCTTACAAAGAC 179
DB |||||

QY 446 TAGTTCTCTCAAAACTACATGAAACCTTCGTTACCCATCTCGCTCGTGGCTTATTTA 505
DB |||||

QY 180 TAGATCTCTCAAAACTACATGAAACCTTCGTTACCCATCTCGCTCGTGGCTTATTTA 239
DB |||||


```

Db      831 AGCCATACTACATACACAAATTCCTCATGATCAAGGTGTAACCTCTCCACAC-A 889
QY      781 ATAGTCTGCTACCTCAGCAATATTTT 809
Db      890 ATAGTCTGCTACCTCAGCAATATTTT 918

RESULT 20
LOCUS   AUI38405              723 bp    mRNA    linear    EST 02-AUG-2002
DEFINITION AUI38405 PLACE1 Homo sapiens cDNA clone PLACE1008489 5', mRNA
sequence.
ACCESSION AUI38405
VERSION   AUI38405.1 GI:10999926
KEYWORDS EST. Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 723)
AUTHORS   Oka,T., Nishikawa,T., Suzuki,Y., Iehii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE     HRI human cDNA project
JOURNAL   Unpublished (2000)
COMMENT   Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute
          1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986
          Email: genomics@hri.co.jp
          HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
          Research Institute; cDNA library construction: Department of
          Virology, Institute of Medical Science, University of Tokyo, and
          Helix Research Institute.
FEATURES             Location/Qualifiers
     source           1..723
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1008489"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"

ORIGIN
Query Match      43.7%; Score 647; DB 1; Length 723;
Best Local Similarity 94.6%; Pred. No. 4.9e-180;
Matches 679; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY      565 TGCCTCCCCCTGCACCTTCAGCCCATACATTTCAATCCCTGTCTCTGACCAATGGAACAAC 624
Db      3   TGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAAC 62
QY      625 TTCAGCAGAAATAAACACACACTTCGTTTTAGTAGGACCTCTGTTTCCAATCTGAA 684
Db      63   TTCAGCAGAAATAAACACACACTTCGTTTTAGTAGGACCTCTGTTTCCAATCTGAA 122
QY      685 ATACCCCATACCTCAAACTCACCTGTGTAAATTTAGCAATACTATAGACACAACACAGC 744
Db      123 ATACCCCATACCTCAAACTCACCTGTGTAAATTTAGCAATACTATAGACACAACAC 182
QY      745 TCCCAATGCATCAGGTGGGTAAACACCTCCCAACAGAAATAGTCTGCCTACCCCTCAGGAATA 804
Db      183 TCCCAATGCATCAGGTGGGTAACTCTCTCCCAACAAATAGTCTGCCTACCCCTCAGGAATA 242
QY      805 TTTTGTGTGTGGTACCTCAGCCCTATCATTTGTTTGAATGGCTCTTTCGAATCTATGTGC 864
Db      243 TTTTGTGTGTGGTACCTCAGCCCTATCGTTGTTTGAATGGCTCTTTCGAATCTATGTGC 302
QY      865 TTCCTCTCATTTCTTAGTGTCCCTCATGACCTATGACCTGACCAAGATTTATACATCAT 924

```

```

Db      303 TTCCTCTCATTTCTTAGTGCCCTTATGACCATCTACATGAAACAAGATTTATACAGTTAT 362
QY      925 GTCTGACCTTAAGCCCCCAACAAAAGAGTACCCATTTCTTCTTTTGTATACAGACAGGA 984
Db      363 GTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTTCTTCTTTTGTATAGGACAGGA 422
QY      985 GTGCTAGGCAGACTAGGTACTGGCATTTGGCAGTATCACAACTCTCACTAGTTCTACTAC 1044
Db      423 GTGCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATCACAACTCTCACTAGTTCTACTAC 482
QY      1045 ARACTATCTCAAGAAATAAATGTTGACATGGAACAGTCACTGACTCCCTGGTCACTCTTG 1104
Db      483 AACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGCCACTCCCTGGTCACTCTTG 542
QY      1105 CAAAGATCAAACTTAACCTCCCTAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTG 1164
Db      543 CAAAGATNAACTTAACCTCCCTAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTG 602
QY      1165 CTAACGCCAAAGAGGGGGAACCTGTTATTTTATAGGAGAAAGCGCTGTTATATGTT 1224
Db      603 CTAACCGCTGAAAGANGGGGAACCTGTTTAAATTTTAGGGGAAGAATGCTCTTATTANGTT 662
QY      1225 AATCAATCCAGATTTCTCACT-CAGAAAGTTAAAGAAATTCGAGATCGAATACAATCT 1281
Db      663 AATCAATCCGGGATCGTCACTGGGAGAAAGTTAAGAAATTCGAGATCGAATACAACGT 720

RESULT 21
LOCUS   AQ261133              631 bp    DNA    linear    GSS 24-OCT-1998
DEFINITION CITBI-E1-2506C15.TF CITBI-E1 Homo sapiens genomic clone 2506C15,
genomic survey sequence.
ACCESSION AQ261133
VERSION   AQ261133.1 GI:3787657
KEYWORDS GSS.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 631)
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
          Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
          Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
          Venter,J.C.
          Use of a random human BAC End Sequence Database for Sequence-Ready
          Map Building
          Unpublished (1998)
          Other_GSSes: CITBI-E1-2506C15.TR
          Contact: Mark Adams
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0200
          Fax: 301 838 0208
          Email: mdadams@tigr.org
          Clones are available from Research Genetics (info@resgen.com). BAC
          end search page:
          http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
          Seq primer: M13-21
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..631
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2506C15"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBelosAC11; Site_1: EcoRI; Site_2: EcoRI;
                     Caltech Human BAC Library D"

ORIGIN
Query Match      42.5%; Score 629.4; DB 8; Length 631;

```

Best Local Similarity 99.8%; Pred. No. 7.8e-175;
Matches 630; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 378 AAACAAGTAAGGAAGCAATCTCCAACTGACCCGGGACATACAGCCCTAGCCCTA 437
Db 1 AAACAAGTAAGGAAGCAATCTCCAACTGACCCGGGACATACAGCCCTAGCCCTA 60

QY 438 CAAGGACTAGTCTCTCAAACTACATGAACCCCTCGTATCCCATCTCGCTGCTGAG 497
Db 61 CAAGGACTAGTCTCTCAAACTACATGAACCCCTCGTATCCCATCTCGCTGCTGAG 120

QY 498 CTAATTTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCCCAAAACCCCTACTAACTG 557
Db 121 CTAATTTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCCCAAAACCCCTACTAACTG 180

QY 558 TTGGATGTCCTCCCTCGCTCACTCAGGACATACATTTCAATCCCTGCTTGAACAATG 617
Db 181 TTGGATGTCCTCCCTCGCTCACTCAGGACATACATTTCAATCCCTGCTTGAACAATG 240

QY 618 GAACAACCTTCAGCAGAGAATAAACAACCACTTCCGTTTTAGTAGGACCTCTGTTTCCA 677
Db 241 GAACAACCTTCAGCAGAGAATAAACAACCACTTCCGTTTTAGTAGGACCTCTGTTTCCA 300

QY 678 TCTGGAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATATAGACAC 737
Db 301 TCTGGAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATACATATAGACAC 360

QY 738 AACAGCTCCCAATGATCAGTGGGTAAACCTCCACACAGATAGTCTGCTACCTC 797
Db 361 AACAGCTCCCAATGATCAGTGGGTAAACCTCCACACAGATAGTCTGCTACCTC 420

QY 798 AGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTGATGGCTCTTCAGAATC 857
Db 421 AGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTGATGGCTCTTCAGAATC 480

QY 858 TATGCTCTCTCTCATTTAGTGCCCTCATGACCATCTACACTGAACAGATTTATA 917
Db 481 TATGCTCTCTCTCATTTAGTGCCCTCATGACCATCTACACTGAACAGATTTATA 540

QY 918 CAATCATGTCGTACTAAGCCCAACAAAGAGATACCATCTCTCTTTGTTATCAG 977
Db 541 CAATCATGTCGTACTAAGCCCAACAAAGAGATACCATCTCTCTTTGTTATCAG 600

QY 978 ASCAGAGTCTAGGACAGACTAGGTACTGGC 1008
Db 601 AGCAAGAGTCTAGGACAGACTAGGTACTGGC 631

RESULT 22
AUI38097
LOCUS AUI38097 PLACE1 Homo sapiens cDNA clone PLACE1007839 5', mRNA
DEFINITION AUI38097 702 bp mRNA linear EST 02-AUG-2002
sequence.
ACCESSION AUI38097
VERSION AUI38097.1 GI:10999618
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 702)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
Location/Qualifiers
1. .702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1007839"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Vector: pME18SFL3"

ORIGIN
Query Match 42.4%; Score 628.6; DB 1; Length 702;
Best Local Similarity 94.6%; Pred. No. 1.4e-174; Indels 1; Gaps 1;
Matches 660; Conservative 0; Mismatches 37;

QY 633 AGAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATCTGGAATAAACCCA 692
Db 1 AGAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATCTGGAATAAACCCA 60

QY 693 TACCTCAAACTCACCCTGTGTAAAAATTAGCAATACATATAGACACCAACGCTCCCAATG 752
Db 61 TACCTCAAACTCACCCTGTGTAAAAATTAGCAATACATATAGACACCAACGCTCCCAATG 120

QY 753 CATCAGTGGGTAACTCCACACGAAATAGTCTGCTACCTCAGGAATATTTTGT 812
Db 121 CATCAGTGGGTAACTCCACACGAAATAGTCTGCTACCTCAGGAATATTTTGT 180

QY 813 CTGTGGTACCTCAGCTATCATTTGTTGAATGGCTCTTCAGAACTCTATGCTCTCTC 872
Db 181 CTGTGGTACCTCAGCTATCATTTGTTGAATGGCTCTTCAGAACTCTATGCTCTCTC 240

QY 873 ATTCTTAGTGCCCTATGACCATCTACCTGAACAGATTTATACATCATGCTGCTACC 932
Db 241 ATTCTTAGTGCCCTATGACCATCTACCTGAACAGATTTATACATCATGCTGCTACC 300

QY 933 TAAGCCCCCAACAAAGAGTACCCATCTCTCTTTGTTATCAGACGAGGTGCTAGG 992
Db 301 TAAGCCCCCAACAAAGAGTACCCATCTCTCTTTGTTATCAGACGAGGTGCTAGG 360

QY 993 CAGACTAGTACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTATC 1052
Db 361 TGCAGTACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTATC 420

QY 1053 TCAAGAAATAAATGGTGACATGGAAACAGGTCACTGCTCCCTGGTCACTTGCAGATCA 1112
Db 421 TCAAGAAATAAATGGTGACATGGAAACAGGTCACTGCTCCCTGGTCACTTGCAGATCA 480

QY 1113 ACTTAACCTCCCTAGCAGCAGTAGTCCCTTCAAAATCGAAGAGCTTTAGACTTGTAAACGC 1172
Db 481 ACTTAACCTCCCTAGCAGCAGTAGTCCCTTCAAAATCGAAGAGCTTTAGACTTGTAAACGC 540

QY 1173 CAAAGAGGGGAACTGTTTATTTTATAGGAGAGAACGCTGTTATTTATTAATCAATC 1232
Db 541 TGAAGAGGGGAACTGTTTATTTTATAGGAGAGAACGCTGTTATTTATTAATCAATC 600

QY 1233 CAGAATTGTCTCAGAGAAAGTTAAAGAAATTCGAGATCGAATACATATAGTAG-AGCAGAGG 1291
Db 601 CGGAATCTCCTCAGAGAAAGTTAAAGAAATTCGAGATCGAATACATATAGTAG-AGCAGAGG 660

QY 1292 AGCTTCAAAACACCGAAGCTGGGCTCCCTCAGCCAA 1329
Db 661 AGCTTCAAAACACCGAAGCTGGGCTCCCTCAGCCAA 698

RESULT 23
BX409328
LOCUS BX409328 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE0131120
DEFINITION 5-PRIME, mRNA sequence.

BX409328
 ACCESSION BX409328.2 GI:46932867
 VERSION
 EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 924)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30652931.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0BAE0122F02_AE01071_i&c=4215.r

Location/Qualifiers
 1..924
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE013V120"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

Query Match 42.1%; Score 623.8; DB 5; Length 924;
 Best Local Similarity 85.5%; Pred. No. 4e-173;
 Matches 748; Conservative 0; Mismatches 120; Indels 7; Gaps 5;

QY 1 ATGGCCCTCCCTATCATACATTTTCTTTACTGTCTCTTACCCCTTTGGCTCTCACT 60
 DB 51 ATGGCCCTCCCTANATATTTTCTTTACTGTCTTTTACCCCTTTTCACTCTCACT 110
 QY 61 GCACCCCTCCATGCTCTGTACACCAAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
 DB 111 GCACCCCTCCATGCTCTGTATGACCAAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 170
 QY 121 ACAGCGGCTCTCTGGAAATTTGATGCCCCATCATATAGGAGTTTATCTAAGGGAATCTC 180
 DB 171 ATGACAGGTCCTCGGAATATTTGATGCCCCATCGTATAGGAGTTTCTAAGGGAATCTC 230
 QY 181 ACCTTCACTGCCACACCATATATGCCCCGCAACTGTCTAATCTGACCTCTTTGCATG 240
 DB 231 ACCTTCACTGCCACACCATATATGCCCCGCAACTGTCTAATCTGACCTCTTTGCATG 290
 QY 241 CATGCAAACTACTCATTTTGGACAGGAAATGATTAATCTAGTTTCTCTGGAGGACTT 300
 DB 291 CATGCAAACTACTCATTTTGGACAGGAAATGATTAATCTAGTTTCTCTGGAGGACTT 350
 QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCAAGTATGTTCTGATGGGGTGGAAAT 360
 DB 351 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCAAGTATGTTCTGATGGGGTGGAGTT 410
 QY 361 CAAGGTGAGGCAAGAGAAAACAAGTAAAGGAGCAATCTCCCAACTGACCCCGGACAT 420
 DB 411 CAAGGTGAGGCAAGAGAAAACAAGTAAAGGAGCAATCTCCCAACTGACCCCGGACAT 470
 QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCTCAAAACTACATCAAAACCCCTCCGTACC 480

DB 471 GGCACCTCTAGCCCTCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACC 530
 QY 481 CATACTCGCCTGCTGAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
 DB 531 CATACTCGCCTGCTGAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 590
 QY 541 CAAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
 DB 591 CAAAACCCCTACNNTACTGTGGATATGCTCCCTCCCTGACTTANNGCATATNGTT--CATC 648
 QY 601 CTTGTTCTTGAACAATGGAACAACTTCAGCACAGAAATAACACCACTTCCTGTTTAGTA 660
 DB 649 CTTGTTCTTGAACAATGGAACAACTTCAGCACAGAAATAACACCACTTCCTGTTTAGTA 707
 QY 661 GGACCTCTTGTTCCTCAATCTGGAATAACCCATCTCAAACTCACTGTGTAAATTTT 720
 DB 708 GGACCTCTTGTTCCTCAATCTGGAATAACCCATCTCAAACTCACTGTGTAAATTTT 764
 QY 721 AGCAATACTATAGACAAACCACTCCCAATGCATCAGGTGGGTAAACACTCCCAACGA 780
 DB 765 AGCAATACTATAGACAAACCACTCCCAATGCATCAGGTGGGTAAACACTCCCAACGA 824
 QY 781 ATAGTCTGCTTACCTCCCTCAGAAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 840
 DB 825 ATAGTCTGCTTACCTCCCTCAGAA-ATTTTGTGGCTGTGGACCTTAACCTTTGTTGTTG 883
 QY 841 AATGGCTCTTCAGAAATCTATGCTTCCTCTCAAT 875
 DB 884 AAGGCTCTTAAAGAAACTATAGCCCTCTCTCAAT 918

RESULT 24
 BE734284 714 bp mRNA linear EST 15-SEP-2000
 LOCUS 601565487F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3840572 5',
 DEFINITION mRNA sequence.
 ACCESSION BE734284
 VERSION BE734284.1 GI:10148276
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 714)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CH530 row: m column: 21
 High quality sequence stop: 712.
 Location/Qualifiers
 1..714
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3840572"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 21"
 /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University

FEATURES
 source

of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN	Query Match		41.2%; Score 609.8; DB 2; Length 714;		DB 2; Length 714;	
	Best Local Similarity	Matches	67%; Conservative	0; Mismatches	37; Indels	4; Gaps
QY	723	CAATCTATAGACACACACAGCTCCCAATGCATCAGTGGGTAACACCTCCACACGAAT	782			
Db	1	CAATCTATACATACAAACCACTCCCAATGCATCAGTGGGTAACCTCCCTCCACACAAAT	60			
QY	783	AGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCTATCAATGTTGAA	842			
Db	61	AGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACTCTCAGCTATC-GTGTGAA	119			
QY	843	TGGCTCTTCAGAACTATGCTCTCTCTCAATCTTAGTGCCCTATGACCATCTACAC	902			
Db	120	TGGCTCTTCAGAACTATGCTCTCTCTCAATCTTAGTGCCCTATGACCATCTACAC	179			
QY	903	TGAACAGATTTATACATCATGCTGCTACCTAAGCCCAACAAAGAGTACCCATTCT	962			
Db	180	TGAACAGATTTATACATCATGCTGCTACCTAAGCCCAACAAAGAGTACCCATTCT	239			
QY	963	TCCTTTTGTATCAGACAGGAGTGTAGGACAGCTAGGTACTGCGATTGGCAGTATCAC	1022			
Db	240	TCC-TTTGTATAGGACAGGAGTGTAGGTGCACTAGGTACTGGCA-TGGCGGTATCAC	297			
QY	1023	AACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGGTGACATGGAACAGT	1082			
Db	298	AACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGGTGACATGGAACGGGT	357			
QY	1083	CAGTACTCTCTGTCACCTTCGACATCACTTAACCTCCCTAGCAGAGTACTCTCA	1142			
Db	358	CGCCGACTCCCTGGTCACCTTCGACATCACTTAACCTCCCTAGCAGAGTACTCTCA	417			
QY	1143	AAATCGAAGAGCTTTAGACTCTGCTAACCGCAAAAGAGGGGAACTGTTATTTTAGG	1202			
Db	418	AAATCGAAGAGCTTTAGACTCTGCTAACCGCTGAAGAGGGGAACTGTTATTTTAGG	477			
QY	1203	AGAAGAACGCTGTTATTTATGTTAAATCAATCCAGAATTTGCTAGAGAAATTTAAAGAAAT	1262			
Db	478	GAAGAAATGCTGTTATTTATGTTAAATCAATCCAGAATTCGCTAGAGAAATTTAAAGAAAT	537			
QY	1263	TCGAGATCGAATACATGATGAGCAGAGGAGCTTCAAAACACCGAACCTGGGGCTCT	1322			
Db	538	TCGAGATCGAATACATGATGAGCAGAGGAGCTTCAAAACACCTGGGGCTCTCT	597			
QY	1323	CAGCAATGGATGGCTCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTGT	1382			
Db	598	CAGCAATGGATGGCTCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTGT	657			
QY	1383	ACTCTCTTTGGACCTGATCTTTAACTCTCTGTTAAGTTTGTCTCTTCAGAAAT	1439			
Db	658	ACTCTCTTTGGACCTGATCTTTAA-CTCCTTGTAACTTTGTCTCTTCAGAAAT	713			

RESULT 25	BX430050/c	658 bp	mRNA	linear	EST 04-MAY-2004
LOCUS	BX430050	Homo sapiens PLACENTA	Homo sapiens	cdna	clone CS0DE012VJ24
DEFINITION	3-PRIME, mRNA	sequence.			
ACCESSION	BX430050				
VERSION	BX430050.2	GI:47003301			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
	Unpublished (2001)				

COMMENT

On May 15, 2003 this sequence version replaced gi:30776872.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0BAB002ZD11_CS00181_1&c=4215.r

FEATURES

source
1..658
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012VJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

	Query Match	39.2%; Score 580.4; DB 5; Length 658;		DB 5; Length 658;
		Best Local Similarity	Pred. No. 2.7e-160;	
	Matches	616; Conservative	0; Mismatches	41; Indels
QY	505	AATACACCCCTCACTCGCTCCATGAGTCTCAGCCCAA-ACCCTACTACTGTGGAT	563	
Db	658	AAAACCATCCACATTTGGATACCAATAGGTTTTTGTACAAATACCCCTACTTAATGTGAT	599	
QY	564	GTGCTCCCTCGACATTCAGGCCATCATTTCAATCCCTGCTGAACCAATGGAACA	623	
Db	598	ATGCTCCCTCGACATTCAGGCCATCATTTCAATCCCTGCTGAACCAATGGAACA	539	
QY	624	CTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTAGGACCTCTTGTTCCTCAATCTCGA	683	
Db	538	CTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTAGGACCTCTTGTTCCTCAATCTCGA	479	
QY	684	AATACCCATACCTCAACCTCAGCTGTGTAATAATTTAGCAATACTATAGACACACAG	743	
Db	478	AATACCCATACCTCAACCTCAGCTGTGTAATAATTTAGCAATACTATAGACACACCA	419	
QY	744	CTCCCAATGCATCAGGTGGGTAACAACCTCCACACGAATAGTCTGCCTTACCTCAGGAAT	803	
Db	418	CTCCCAATGCATCAGGTGGGTAACCTCCACACGAATAGTCTGCCTTACCTCAGGAAT	359	
QY	804	ATTTTGTCTGTGGTACCTCAGCTATCATTTGTTGAATGGCTCTTCAGAAATCTATGTG	863	
Db	358	ATTTTGTCTGTGGTACCTCAGCTATCATTTGTTGAATGGCTCTTCAGAAATCTATGTG	299	
QY	864	CTTCTCTCTCTTAGTGGCCCTTACCATCTAGCACTCAACAGATTTATACAATCA	923	
Db	298	CTTCTCTCTCTTAGTGGCCCTTACCATCTAGCACTCAACAGATTTATACAATCA	239	
QY	924	TGTCATATCAATGAGCCCGCAACAAAGAGTACCCATCTCTCTTTGTTATGAGGACAGG	983	
Db	238	TGTCATATCAATGAGCCCGCAACAAAGAGTACCCATCTCTCTTTGTTATGAGGACAGG	179	
QY	984	AGTGTAGGACAGCTAGGTACTGGCATATGACAACTCTACTAGTCTACTA	1043	
Db	178	AGTGTAGGACAGCTAGGTACTGGCATATGACAACTCTACTAGTCTACTA	119	
QY	1044	CAAACTATCTCAAGAAATAATGTCATGGAACAGTCTACTCTCCCTGGTCACTT	1103	
Db	118	CAAACTATCTCAAGAAATAATGTCATGGAACAGTCTACTCTCCCTGGTCACTT	59	

ORIGIN		Query Match	36.3%	Score 537.6	DB 5	Length 814
		Best Local Similarity	85.3%	Pred. No. 1.4e-147		
		Matches	671	Conservative	0	Mismatches 105; Indels 11; Gaps 6
Qy	379	AAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACATAGCACCCCTAGCCCTAC	438			
Db	814	AAAGCTGGACACTAGAAATTTCTCAACCCACCGGAGCAAGCACCCCTATACCCCTTA	755			
Qy	439	AAAGGACTAGTTCTCTCAAAACTACATGAAGAACCCCTCCGTACCCCATACTCGCCT	497			
Db	754	AAAGGTATAAACTGCCAAAATAACACTAAACCCCGGGGCCAAATATGCCTGGGTAA	695			
Qy	498	CTATTTAAATACACCCCTCACTCGGCTCCATGAGGCTCTCAGCCCAAAACCCCTACTAACTG	557			
Db	694	CCAATTTAAACACCAACCTCA--TGCGGCCCATAGGCTCGGCCCAAAACCCCTACTAACTG	637			
Qy	558	TTGATGTGCTCCCGCTGCACTT--CAGGCCATACATTTCAATCCCTGTTCTCTGAA-C	613			
Db	636	TTGATATGCTTCCCGCTGNACTTCAAGGCCAAATGTTCAAAATCCCTGACTCTGNAACC	577			
Qy	614	AATGGAAACAATT---CAGCACAGAAATAAACACCACTTCCGTTTTTASTAGGACCTCTTG	670			
Db	576	AATGGANCAACTTTCAGCACAGAAAAATAAACACCAACTTCCGTTTTTATAGTAGGACCTCTTG	517			
Qy	671	TTTCCAACTCGGAAT--AACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATACT	729			
Db	516	TTTCCAATCTGGAAATAAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCAATACT	457			
Qy	730	ATAGACAACAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCCAACAAGATAGTCGC	789			
Db	456	ACATACACAACCAACTCCCAATGATCAGGTGGGTAACTCCTCCCAACAATAGTCTGC	397			
Qy	790	CTACCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCATATCATTTGTTGAATGGCTCT	849			
Db	396	CTACCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCATATCGTTGTTGAATGGCTCT	337			
Qy	850	TCGAAATCTATGCTTCCCTCATCTTAGTGCCCTCTATGACATCTACACTGAACAA	909			
Db	336	TCGAAATCTATGCTTCCCTCATCTTAGTGCCCTCTATGACATCTACACTGAACAA	277			
Qy	910	GATTTATACAATCATGTGTAACCTTAAGCCCAACAACAAAGAGTACCCATCTTCCTTTT	969			
Db	276	GATTTATACAATATATGTATATCTTAAGCCCGCAACAACAAAGAGTACCCATCTTCCTTTT	217			
Qy	970	GTTATCAGCAGGAGTGCTAGGCAGACTAGGTACTGGCAATGGCAGTATCAACCTCT	1029			
Db	216	GTTATAGCAGGAGTGCTAGGTGCACTAGGTACTGGCAATGGCGGTATCAACACCTCT	157			
Qy	1030	ACTCAGTTCTACTACAAACTATCTCAAGAATAAATGGTGACATGGAAACAGGTCACTGAC	1089			
Db	156	ACTCAGTTCTACTACAAACTATCTCAAGAATAAATGGTGACATGGAAACAGGTCACTGAC	97			
Qy	1090	TCCCTGTCTACCTTGGCAAGATCAACTTAACCTCCCTAGCAGAGTAGTCTTCAAAATCGA	1149			
Db	96	TCCCTGTCTACCTTGGCAAGATCAACTTAACCTCCCTAGCAGAGTAGTCTTCAAAATCGA	37			
Qy	1150	AGAGCTT 1156				
Db	36	AGAGCTT 30				
RESULT 28						
BE019603		LOCUS				
DEFINITION		BE019603 586 bp mRNA linear EST 06-JUN-2000 ba84f03.y1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:2907101 5' similar to TR:Q95244 Q95244 ENVELOPE PROTEIN ; contains Alu repetitive element?, mRNA sequence.				
ACCESSION		BE019603				
VERSION		BE019603.1				
KEYWORDS		EST.				
SOURCE		Homo sapiens (human)				

1183 GGAACCTGTTTATTTTAGGAGAAACGCTGTTATTTAATCAATCCAGAAATGTC 1242
 481 GGAACCTGTTTATTTTAGGAGAAATGCTGTTATTTAATCAATCCGAAATCGTC 540
 1243 ACTGAGAAATGTTAAAGAAATTCGAGATCGAATACAAATGAGAGCAG 1288
 541 ACTGAGAAATGTTAAAGAAATTCGAGATCGAATACAAATGAGAGCAG 586

RESULT 29
 BG572445 790 bp mRNA linear EST 10-APR-2001
 LOCUS 602593450F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4720846 5',
 DEFINITION mRNA sequence.
 ACCESSION BG572445
 VERSION BG572445.1 GI:13580098
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 790)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-f@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 plate: LICM1576 row: c column: 23
 High quality sequence stop: 624.

FEATURES
 Location/Qualifiers
 1..790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4720846"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NIH_MGC_79"
 /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
 Site 1: Sfil (ggccgctcgcc); Site 2: Sfil
 (ggccattatggc); 5' and 3' adaptors were used in cloning
 as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3',
 and 3' adaptor sequence:
 5'-ATTCTAGGCGCGGGCGGCACATG-DT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.3
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 36.1%; Score 535.2; DB 4; Length 790;
 Best Local Similarity 90.5%; Pred. No. 7.3e-147;
 Matches 650; Conservative 0; Mismatches 58; Indels 10; Gaps 7;

742 AGCTCCCAATGATCAGTGGTGAACACCTCCACAGAAATAGTCTGCCTACCTCAGGA 801
 1 AACTCCCAATGATCAGTGGTGAACCTCCACAGAAATAGTCTGCCTACCTCAGGA 60
 802 ATATTTTGTCTGCTGTGCTACCTACCTATCATTTGTTGATGCTCTTCAGAACTATG 861
 61 ATATTTTGTCTGCTGTGCTACCTACCTATCATTTGTTGATGCTCTTCAGAACTATG 120
 862 TGCTTCCTCTCATTTCTTAGTGCCTCCCTACCACTACCACTGACAAAGATTTTACAAT 921
 121 TGCTTCCTCTCATTTCTTAGTGCCTCCCTACCACTACCACTGACAAAGATTTTACAAT 180

922 CATGTGCTACTTAAGCCCAACAAAGAGTAGTACCATTCTCTCTTTGTTTATCAGAGCA 981
 181 TATGTATCATATCTAAGCCCGCAACAAAGAGTAGTACCATTCTCTCTTTGTTTATCAGAGCA 240
 982 GGAAGTCTAGGCAGACTAGGTACTGCGCATTTGGCAGTATCACAACCTCTACTCAGTCTCTAC 1041
 241 GGAAGTCTAGGTGCACTAGGTACTGCGCATTTGGCAGTATCACAACCTCTACTCAGTCTCTAC 300
 1042 TACAAACTATCTCAAGAAATAAATGGTGCATGGAACAGGTCACTGACCTCCCTGGTCAACC 1101
 301 TACAAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGCGACTCCCTGGTCAACC 360
 1102 TTGCAAGATCAACTTAATCTCCCTAGCAGCAGTAGTCTCTCA-AAATCGAAGAGCTTTAGA 1160
 361 TTGCAAGATCAACTTAATCTCCCTAGCAGCAGTAGTCTCTCACAAATCGAAGAGCTTTAGA 420
 1161 CTTGCTTAACCGCCAAAGAGGGGGAACCTGTTTATTTTAGGAGAAAGAACGCTGTTTATTA 1220
 421 CTTGCTTAACCGCTGAAGAGGGGGAACCTGTTTATTTTAGGAGAAAGAACGCTGTTTATTA 480
 1221 TGTTAATCAATCCAGAAATTTGTCACTGAGAAAGTTA-AAGAAATTCGAGATCG--AATACA 1277
 481 TGTTAATCAATCGGNAATCGTCACTGAGAAAGTTAACAAGAAATTCGAGATCGCAATCAC 540
 1278 ATGTAGAGCAGAGGAGCTTTCAAAACACCGAAGCTGGGGCCCTCTCAGCCAATGGATGCC 1337
 541 ACGTAGAGCAGAGGAGCTTTCAAAACACCTGACCTCGGGGC--TCCTCAGCAATGGATGCC 598
 1338 CTTGGGTTCTCCCTCTTTAGGACCTCTAGC-AGCTCTAATATTTGTTACTCTCTTTGGAC 1396
 599 --TGGATTCTCCCTCTTTAGGACCTCTAGCAAGCTATATCTATTGCTACTCTCTC-TTGGAG 655
 1397 CTTGATCTTTAACTCTCTTTGTTAAGTTTGTCTCTTCCAGAAATGGAAGCTGTTAAAGCT 1454
 656 CTTGATCTTTAACTCTCTTTGTTAAGTTTGTCTCTTCCAGAAATGGAAGCTGTTAAAGCT 713

RESULT 30

BX450495/c 760 bp mRNA linear EST 05-MAY-2004
 LOCUS BX450495 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120
 DEFINITION 3-PRIME, mRNA sequence.
 ACCESSION BX450495
 VERSION BX450495.2 GI:47038997
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 760)
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 22, 2003 this sequence version replaced gi:31018916.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4215.r
 For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AU006ZG02_U0495_1kc=4215.r.
 Location/Qualifiers
 1..760
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE013Y120"
 /tissue_type="PLACENTA"

FEATURES
 source

/clone lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN	Query Match Best Local Similarity 32.3%; Score 478.6; DB 5; Length 760; Matches 604; Conservative 0; Mismatches 114; Indels 7; Gaps 5;
Qy	682 GAAATAAACCCATACCTCAAACTCCTACCTGTGTAAAAATTAGCAATCT-ATAGACACAAC 740
Db	727 GCACACACCAAGCCCTCCCTCGGGGGCAAAATCTCGTATCTCCCTATATCAAC 668
Qy	741 CAGCTCCCAATGCATCAG---GTGGGTAAACACTCCCAACAGAAATAGTCTGCCTACCTC 797
Db	667 CGACCAACCATGCATCAGGTGAGTGGTCCCTCTCTCAAAAATAGTTTGCTATACCTCC 608
Qy	798 AGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTGAATGCTCTTCAGATC 857
Db	607 AGAATATAGTGGTGTGTGAGGCTCTCTCAGCCAAACCGTAGTAGAAGGGCTCTTCAGAAC 548
Qy	858 TATGTGCTTCCTCTCATTTAGTGCCTTATGACCATCTACACTGAAAC-AAAGATTAT 916
Db	547 AATGTGC-TCCCTGCANCTAGAGCCCTTATGACCTGACACTGAACTTAGATTAG 489
Qy	917 ACAATCATGTCGTACTAAGCCCCCAACAAAAGAGTAGTACCCATTTCTCTTTTGTATCA 976
Db	488 CCCATTATGTATATCTAAGCCAGCAACAAAAGAGTCCCATTTCTCTCTTTTGTATAG 429
Qy	977 GAGCAGGAGTGTAGGACACTAGTACTGCGATTGGCAGTATCAAA-CCTCTACTCAG 1035
Db	428 AAGCAGGAGTGTAGGACACTAGGAACCTGGCAATGTGGTATCAAAAGCCTCTACTCAG 369
Qy	1036 TTCTACTACAAATCTCTCAAGAAATAAATGGTGACATGGAACAGTCACTGACTCCCTG 1095
Db	368 TTCTACTACGGAATCTCAAGACTAATATGGGACATGGAACGGGTCCCGCACTCCCTG 309
Qy	1096 GTCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTAGTCCCTTCAAAATCGAAGCT 1155
Db	308 GTCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTAGTCCCTTCAAAATCGAAGCT 249
Qy	1156 TTAGACTTCTAACCCCAAGAGGGGCAACCTGTTATTTTATAGGAGAAAGCGCTGT 1215
Db	248 TTAGACTTGTCTAACCCGCTGAAAGAGGGGGGACCTGTTTATTTTATAGGGGAAGATGCTGT 189
Qy	1216 TATTATGTTAAATCAATCCAGAAATTTGCTACTGAGAAAGTTAAAGAAATTCGAGATCGAATA 1275
Db	188 TACTATGTTAAATCAATCCGGAATCGTACTGAGAAAGTTAAAGAAATTCGAGATCGAATA 129
Qy	1276 CAATGTAGACGAGAGAGCTTCAAAACACCGAACGCTGGGCTCTCTAGCCCAATGGATG 1335
Db	128 CAACGTAGACGAGAGAGCTTCCAAACACTGGACCTGGGGCTCTCTAGCCCAATGGATG 69
Qy	1336 CCCTGGGTTCTCCCTTCTTAGACCTCTAGCAGCTCTAATATTTGTTACTCCTCTTGA 1395
Db	68 CCCTGGATTTCCCTTCTTAGACCTCTAGCAGCTCTAATATTTGTTACTCCTCTTGA 9
Qy	1396 CCCTG 1400
Db	8 CCCTG 4

RESULT 31
CR735592
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

CR735592 554 bp mRNA linear EST 27-AUG-2004
CR735592 NIH MGC 21 Homo sapiens cDNA clone IMAGE3840572 5', mRNA sequence.
IMAGE:3840572 5', mRNA sequence.
CR735592
CR735592.1 GI:51584854
EST.

SOURCE ORGANISM	Homo sapiens (human) Homo sapiens Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE AUTHORS	1 (bases 1 to 554) Ebert,L., Heil,O., Hennig,S., Korn,B., Neubert,P., Partsch,E., Peters,M., Radelof,U. and Schneider,D. I.M.A.G.E. cDNA Clone Collection
TITLE JOURNAL COMMENT	Unpublished (2004) Contact: Inge Arlart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Email: www.rzpd.de RZPD; IMAGp958M21530. RZPDLIB; I.M.A.G.E. cDNA Clone Collection; Contact: Inge Arlart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 100 Fax: +49 30 32639 111 www.rzpd.de This clone is available royalty-free from RZPD; contact RZPD (clone@rzpd.de) for further information. Seq primer: M13r, Primer sequence: TTTCACACAGAAACAGCTAAGAC. Location/Qualifiers 1..554 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGp958M21530 ; IMAGE:3840572" /tissue_type="choriocarcinoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 21" /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
FEATURES source	
ORIGIN	Query Match 32.1%; Score 475.4; DB 7; Length 554; Best Local Similarity 95.0%; Pred. No. 3.6e-129; Matches 491; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy	702 CCTCACCTGTGTAATAATTTAGCAATATCTATAGACACACACAGCTCCCAATGCATCAGGTG 761
Db	38 CCTCACCTGTGTAATAATTTAGCAATATCTATAGACACACACAGCTCCCAATGCATCAGGTG 97
Qy	762 GGTAAACCTCCACACGAAATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGGTAC 821
Db	98 GGTAACTCTCTCCACACAAATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGGTAC 157
Qy	822 CTCAGCCTTATCATTTGTAATGGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTAGT 881
Db	158 CTCAGCCTTATCGTTTGTGTAATGGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTAGT 217
Qy	882 GCCCCTATGACCATCTACACTGAAACAGATTTATACAAATCATGTGCTACCTACGCCCA 941
Db	218 GCCCCTATGACCATCTACACTGAAACAGATTTATACAGATTATGTGCATATCTAAGCCCG 277
Qy	942 CAACAAAGAGTAGCCATTTCTCTTTGTTTATCAGACGAGGTGCTAGGCAGACTAGG 1001
Db	278 CAACAAAGAGTAGCCATTTCTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGG 337
Qy	1002 TACTGGCATTGGCAGTATCAAACTCTTACTCAGTTCTTACTACAACTATCTCAAGAAAT 1061
Db	338 TACTGGCATTGGCGGTATCAAACTCTTACTCAGTTCTTACTACAACTATCTCAAGAACT 397
Qy	1062 AAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAATC 1121

Db 847 TCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGTCAACCGCTAAAGA 788

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0BAA015ZF07_CS01399_1&c=4215.r

FEATURES

source
1. .890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1022YJ18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 28.1%; Score 416.4; DB 5; Length 890;
Best Local Similarity 83.7%; Pred. No. 1.3e-111;
Matches 543; Conservative 0; Mismatches 97; Indels 9; Gaps 6;
QY 1 ATGGCCCTCCCTTATCATCTTTCTTTACTGTTCTTTACCCCTTTTGGCTCTCACT 60
DB 215 ATGGCCCTCCCTTATCATCTTTCTTTACTGTTCTTTACCCCTTTTGGCTCTCACT 274
QY 61 GCACCCCTCCCTGCTGTACCAACAGTAGTCTCCCTTACCAAGAGTTTCTTATGAAGA 120
DB 275 GCACCCCTCCCTGCTGTACCAACAGTAGTCTCCCTTACCAAGAGTTTCTTATGAAGA 334
QY 121 ACGGCGCTTCTGGAAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 180
DB 335 ATGCAGCGTCCCGAAATATTTGATGCCCATCATATAGGAGTTTCTTCTAAGGGAATCC 394
QY 181 ACCTTCACTGCCACACCCATATGCCCGGCAACTGCTATTAATCTGCGCACTTTTGCATG 240
DB 395 ACCTTCACTGCCACACCCATATGCCCGGCAACTGCTATTAATCTGCGCACTTTTGCATG 454
QY 241 CATGCAATCTCATTTATGGACAGGGAATATTAATCTAGTTTCTTGGAGGACTT 300
DB 455 CATGCAATCTCATTTATGGACAGGGAATATTAATCTAGTTTCTTGGAGGACTT 514
QY 301 GGAGCCACTGCTGTGGAGTCTTCACTCCCATACAGTATGCTGTATGGGGTGGAAAT 360
DB 515 GGAGTCACTGCTGTGGAGTCTTCACTCCCATACAGTATGCTGTATGGGGTGGAGTT 574
QY 361 CAAGGTGAGCAAGAGAAACAAAGTAAGAGCAATCTCCCACTGACCCGGGACAT 420
DB 575 CAAGTACAGCAAGAGAAACAAAGTAAGAGCAATCTTCCCACTTACCCCGGTACAT 634
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTTACATGAACCCCTCCGTACC 480
DB 635 GGCACTTTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTTACATGAACCCCTCCGTACC 692
QY 481 CATACCTGCTGTGGAGTCTTCACTCCCATACAGTATGCTGTATGGGGTGGAGTT 540
DB 693 CATACT-GGCTGGTAAAGCTTATTAATACCCACC--TACTGGCCCTCCATGAGGCTTGGGCC 749
QY 541 CAAGAACCCTACTACTGTTGATGCTGCTCCCTGCACTTTCAGGCCATACATTTCAATC 600
DB 750 AAAAACC--TACTAATGGTGGAAATGCTTCCCGGACATATTAGGCCATAGTTT--AAAT 805
QY 601 CCTGTTCTGAACAATGAACAACCTTCAAGCAAGAAATTAACACCACTT 649
DB 806 CCCTGGCCCGAAACAGGGAACAACCTTTGCCAAGAAATAAACCCCTCTCT 854

RESULT 35
BX337769/c
LOCUS
DEFINITION
BX337769 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1051YJ13 3-PRIME, mRNA sequence.
ACCESSION
BX337769
VERSION
BX337769.2 GI:46272079

QY 1 ATGGCCCTCCCTTATCATCTTTCTTTACTGTTCTTTACCCCTTTTGGCTCTCACT 60
DB 209 ATGGCCCTCCCTTATCATCTTTCTTTACTGTTCTTTTACCCCTTTTGGCTCTCACT 268
QY 61 GCACCCCTCCCTGCTGTACCAACAGTAGTCTCCCTTACCAAGAGTTTCTTATGAAGA 120
DB 269 GCACCCCTCCCTGCTGTATGACCACTAGTCTCCCTTACCAAGAGTTTCTTATGAAGA 328
QY 121 AGCGGCTTCTGGAAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 180
DB 329 ATGCAGCGTCCCGAAATATTTGATGCCCATCATATAGGAGTTTCTTAAAGGGAATCC 388
QY 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATACTGCGCACTTTTGCATG 240
DB 389 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATACTGCGCACTTTTGCATG 448
QY 241 CATGCAATCTCATTTATGGACAGGGAATATTAATCTAGTTTCTTGGAGGACTT 300
DB 449 CATGCAATCTCATTTATGGACAGGGAATATTAATCTAGTTTCTTGGAGGACTT 508
QY 301 GGAGCCACTGCTGTGGAGTCTTCACTCCCATACAGTATGCTGTATGGGGTGGAAAT 360
DB 509 GGAGTCACTGCTGTGGAGTCTTCACTCCCATACAGTATGCTGTATGGGGTGGAGTT 568
QY 361 CAAGGTGAGCAAGAGAAACAAAGTAAGAGCAATCTCCCACTGACCCGGGACAT 420
DB 569 TAAGATCAGTAAGAGAAACAAAGTAAAGAGTTTCTTCACTTACCCGTTTCTAT 628
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
DB 629 GGTTCCTTTAGTCCCTTACAAAGTCTAGTTGTTGTAAGAACTACATGAACCCCTTGGTCC 688
QY 481 CATACT-CGCTGTGGAGCTATTT-ATAACACCCCTCACTCGGC--TCCATGAGTCTC 536
DB 689 CATACTGCTTCTGGTTAGCTGTTGAATACCAACCCCTAAGTCTTCTTATGAGGTTAT 748
QY 537 AGCCCAAAACCCCTACTAACTGTTGATGTCCTCCCTCGCACTTCAGGCCATACATTT 596
DB 749 GACCCAAACCCCTACTAACTGCGGATATGCTGCTCTGTTGAGTGTGTTGTT 808
QY 597 AATCCCTGTTCTGAAATGAACAACT-TCAGCACAGAAATAAACACCACTT 649
DB 809 AATTTCTGTTCCGTTTACATGTTTCAACTGTTGGTGTAGGAATTAACACCCCTT 862

RESULT 34
BX326300
LOCUS
DEFINITION
BX326300 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1022YJ18 5-PRIME, mRNA sequence.
ACCESSION
BX326300
VERSION
BX326300.2 GI:46278854
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 890)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30342581.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r

DB	518	AAATGGAGCCCAAGATGCAGTCCAA	494
RESULT 36			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match	27.8%;	Score 412.2;	DB 5; Length 1058;
Best Local Similarity	87.9%;	Pred. No. 2.5e-110;	
Matches 457;	Conservative 11;	Mismatches 49;	Indels 3; Gaps 2;
QY	962	TTCTCTTTTGTATCAGAGCAGGAGTCTAGCGACACTAGGTACTGGCATTGGCAGTATCA	1021
DB	1007	TTCTCTTTTGTATCAGAGCAGGAGTCTAGCGACACTAGGTACTGGCATTGGCAGTATCA	951
QY	1022	CAACCTCTACTCAGTCTTACTACAAACTATCTCAAGAAATAAATGTTGATGAAACAGG	1081
DB	950	CAACCTCTACTCAGTCTTACTACAAACTATCTCAAGAAATAAATGTTGATGAAACAGG	891
QY	1082	TCACCTGCTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGCAGTACTGCTTC	1141
DB	890	TCGCGCACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGCAGTACTGCTTC	831
QY	1142	AAATCGAAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTCTGTTATTTTATG	1201
DB	830	AAATCGAAGCTTTAGACTTGTCTAACCGCCCTGAAGAGGKXKAWCYTTTTTTTTT	771
QY	1202	GAGAAGAACCTGTTATTTATTTGTTAAATCAATCCAGAAATTTGCTACTGAGAAAGTTAAAGAA	1261
DB	770	GGGAGAAATGCTGTTACTATGTTAAATCAATCCKAWTKCTCACTGAGAAAGTTAAAGAA	711
QY	1262	TTGAGATCGAATACAAATGTAGACGAGAGCTTCAAAACACCGAACGCTGGGCTCC	1321

Db	710	TTCCAGATCGAATACAACTAGACGACGAGGAGCTTCGAAACACACTGGACCCCTGGGGCCCTCC	651
Qy	1322	TCAGCCAAATGGATGGCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAATATTGT	1381
Db	650	TCAGCCAAATGGATGGCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAATATTGC	591
Qy	1382	TACTCTCTTTTGGACCCCTGTATCTTTTAACTCCTCTTTTAAAGTTTGTCTTCTCCAGAAATG	1441
Db	590	TACTCTCTTTTGGACCCCTGTATCTTTTAACTCCTCTTTTAACTTGTCTTCTTCCAGAAATCG	531
Qy	1442	AACTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA	1481
Db	530	AACTGTAAGCTACAGATGGTCTTACAAATGGAACCCCAAGATGCAATGCA	491
RESULT 37			
LOCUS	BX429316	905 bp	mRNA linear EST 05-MAY-2004
DEFINITION	BX429316 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120	5-PRIME, mRNA sequence.	
ACCESSION	BX429316		
VERSION	1	GI:47035285	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	On May 15, 2003 this sequence version replaced gi:30786832. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS0AT011ZF02_T0974_l&c=4215.r.		
FEATURES			
source	1..905	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DE013Y120"		
	/tissue_type="PLACENTA"		
	/clone_lib="Homo sapiens PLACENTA"		
	/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."		
ORIGIN			
Query Match	27.8%;	Score 411.4;	DB 5; Length 905;
Best Local Similarity	82.5%;	Pred. No. 4.1e-110;	
Matches	572;	Conservative 0;	Mismatches 101; Indels 20; Gaps 8;
Qy	1	ATGGCCCTCCCTTATCATATCTTCTTACTGTCTTCTTACCCCTTTCCGCTCTCACT	60
Db	196	ATGGCCCTCCCTTATCATATCTTCTTACTGTCTTCTTACCCCTTTCCGCTCTCACT	255
Qy	61	GCACCCCTCCATGCTGCTGTATGACCAACAGTAGCTCCCTTACCAAGAGTTTCTTATGAAGA	120
Db	256	GCACCCCTCCATGCTGCTGTATGACCAACAGTAGCTCCCTTACCAAGAGTTTCTTATGAAGA	315
Qy	121	ACCGCGCTTCTTGGAAATATTGATGCCCATCATAGAGTTTATCTTAAGGGAATCC	180

/db xref="taxon:9606"
/clone="CS0DI022YJ18"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 27.4%; Score 406.4; DB 5; Length 564;
Best Local Similarity 84.4%; Pred. No. 1.1e-108;
Matches 467; Conservative 0; Mismatches 84; Indels 2; Gaps 1;
QY 813 CTGTGTACCTGAGCTATCATTTGTTGATGCTCTTCAGATCTATGCTGCTCTCTC 872
DB 564 CTGGGAACCAACCAACAGCATGTTAAGAGACAGTAAAGAAACGAGAGAGCTACAGAC 505
QY 873 ATTCTTAGTCCGCCCTAT--GACATCTACACTGMAACAGATTATACATCATGTCGTA 930
DB 504 ATACTAAGAGAGCTATGAGACAACAGCGCTGGAGCAAGATTATCATGTTATGICATA 445
QY 931 CCTAAGCCCCACAAACAGAGTACCCATTCTCTCTTTCTTTATTCAGAGCAGAGTGCTA 990
DB 444 TTTAAGCCCCGCATCAAAAGAGTACCCATTCTCTCTTTCTTTATAGGAGCAGAGTGCTA 385
QY 991 GGCAGACTAGGTACTGGCATTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTA 1050
DB 384 GGTGCACTAGGTACTGGCATTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTA 325
QY 1051 TCTCAAGAAATAATGTGATGACATGGAACAGTCACTGACCTCCCTGGTCACTTGGCAAGAT 1110
DB 324 TCTCAAGAAATAATGGGACATGGAACGGGTGCCGACCTCCCTGGTCACTTGGCAAGAT 265
QY 1111 CAACCTTAACCTCCCTAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTGTAAAC 1170
DB 264 CAACCTTAACCTCCCTAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTGTAAAC 205
QY 1171 GCGAAAGAGGGGAACCTGTTATTTTAGGAGAGAAACGCTGTTATTTATGTTATCAAA 1230
DB 204 GCTGAAGAGGGGAACCTGTTATTTTAGGAGAGAAAGTCTGTTATTTATGTTATCAAA 145
QY 1231 TCCAGATTGTCTACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAG 1290
DB 144 TCCGGAATCTGCTACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACAACTAGAGCAGAG 85
QY 1291 GAGCTTCAAAACACCGAACGCTGGGCGCTCTCTCAGCCAATGGATGCCCTGGGTTCTCCC 1350
DB 84 GAGCTTCAAAACACCTGGGCGCTCTCTCAGCCAATGGATGCCCTGGGTTCTCCC 25
QY 1351 TTCTTAGGACCTC 1363
DB 24 CTNCTAGGACCCC 12

RESULT 39
BX439636/c
LOCUS
DEFINITION BX439636 Homo sapiens PLACENTA COT 25-NORMALIZED EST 04-MAY-2004
3-PRIME, mRNA sequence.
ACCESSION BX439636
VERSION BX439636.2 GI:47000005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1019)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30771765.
Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DE012DE12NP1&c=4215.r.

FEATURES

Location/Qualifiers
1..1019
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 27.1%; Score 402; DB 5; Length 1019;
Best Local Similarity 90.0%; Pred. No. 2.6e-107;
Matches 434; Conservative 6; Mismatches 41; Indels 1; Gaps 1;
QY 1001 GTATGGCATTTGGCAGTATCAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAA 1060
DB 975 GGWAYKCATTTGGCGTWTMTTAAACAACTCTACTCAGTTCTACTCAAACTATCTCAAGAAC 916
QY 1061 TAAATGGTACATGGACAGTCACTGACCTCCCTGGTCACTT--GCAAGATCAACTTAAC 1119
DB 915 TAAATGGGACATGGAAACGGGTCCGCGACTCCCTGGWCACCTTGGCAAGATCAACTTAAC 856
QY 1120 TCCTTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAAACGCAAAAGA 1179
DB 855 TCCTTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAAACGCTGAAAGA 796
QY 1180 GGGGGAACCTGTTATTTTATTTAGGAGAAACCGCTGTTATTTATTAATCAATCCAGATT 1239
DB 795 GGGGGAACCTGTTATTTTATTTAGGAGAAATTCGAGATCGAATACAACTAGAGCAGAGCTTCAA 1299
QY 1240 GTCACTGAGAAAGCTTAAAGAAATTCGAGATCGAATACAACTAGAGCAGAGCTTCAA 1299
DB 735 GTCACTGAGAAAGCTTAAAGAAATTCGAGATCGAATACAACTAGAGCAGAGCTTCAA 676
QY 1300 AACACCGAACGCTGGGCGCTCTCTCAGCCAATGGATGCCCTGGGTTCTCCCTCTCTTAGGA 1359
DB 675 AACCTGGACCTTGGGCGCTCTCTCAGCCAATGGATGCCCTGGGTTCTCCCTCTCTTAGGA 616
QY 1360 CCTCTAGCAGCTCTAATATTGTTTACTCTCTTTGGACCTGTTATTTTAACTCTCTGTT 1419
DB 615 CCTCTAGCAGCTCTAATATTGTTTACTCTCTTTGGACCTGTTATTTTAACTCTCTGTT 556
QY 1420 AGTTTGTCTCTCCAGAAATTAAGCTTAAGCTACAGATGGTCTTACAAATCGAACCC 1479
DB 555 AACTTTGTCTCTCCAGAAATTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTCC 496
QY 1480 CA 1481
DB 495 AA 494

RESULT 40
BX409035/c
LOCUS
DEFINITION BX409035 Homo sapiens PLACENTA COT 25-NORMALIZED EST 03-MAY-2004
3-PRIME, mRNA sequence.
ACCESSION BX409035

VERSION BX409035.2 GI:46956493
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 484)
Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30766926.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAD005ZE05_AD00422_l&c=4215.r

FEATURES
source Location/Qualifiers
1..484
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 27.1%; Score 401.4; DB 5; Length 484;
Best Local Similarity 91.9%; Pred. No. 3.1e-107;
Matches 445; Conservative 0; Mismatches 37; Indels 2; Gaps 2;
QY 681 GGAAATAACCCATACCTCAAACTCACCCTGTGTAAATTTAGCAATCTATAGACACAC 740
DB 484 GGAAGAAGACCCATCCCTCAAGCCTCACCTGTGTAAATTTAGCAATCTATAGACAC 425
QY 741 CAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGAATAGTCTGCTACCCCTCAGG 800
DB 424 CAACTCACAATGATCAGGTGGGTAACTCTCCACACAATAGTCTGCTACCCACAGA 365
QY 801 AATATTTTTTGTCTGTGGTACCTCAGCCTATACATGTTTGAATGGCTCTTCAGAACTCTAT 860
DB 364 AATATTTTTTGTCTGTGGTACCTCAGCCTATGTTTGAATGGCTCTTCAGAACTCTAT 305
QY 861 GTGCTTCTCTCAATTTCTAGTGGCCCTATGACCTACACTGACAGAGATTTATACAA 920
DB 304 GTGCTTCTCTCAATTTCTAGTGGCCCTATGACCTACACTGACAGAGATTTATACAA 245
QY 921 TCATGTCTGTAACCTAAGCCCAACAAAGAGTACCCTTCTTCTTTTGTATCAGAGC 980
DB 244 TTATGTCATATCTAAGCCCGCAACAAGAGTACCCTTCTTCTTTTGTATAGAGC 185
QY 981 AGGAGTCTAGGAGACTAGGTACTGGCAATTGGCAGTATCAACCTCTACTCAGTTCTA 1040
DB 184 AGGAGTCTAGGAGTCTAGGTACTGGCAATTGGCAGTATCAACCTCTACTCAGTTCTA 125
QY 1041 CTACAACTATCTCAAGAAATTAATGGTACATGGAGACAGGTCACTGACT-CCCTGGTCA 1099
DB 124 CTACAACTATCTCAAGAACTTAATGGGACATGGAGACGGGTGCCGACTCCCTGGTCA 65
QY 1100 CCTTGCAGATCAACTTAACTCCCTAGCAG-CAGTAGTCTCTCAAAATCGAAGAGCTTTA 1158

Db 64 CCTTGCAGATCAACTTAACTCCCTAGCAGCCAGTAGTCCCTCAAAATNCGAAGAGCTTA 5
Qy 1159 GACT 1162
Db 4 GACT 1
Search completed: February 21, 2005, 14:51:11
Job time : 4614.5 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:32:54 ; Search time 5704.48 Seconds
(without alignments)
11288.850 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatgaagagcttttag.....gcaagatccccagtaaca 1329

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

RESULT 1
BD136199
LOCUS
DEFINITION
BD136199 1329 bp DNA linear PAT 18-SEP-2002
Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD136199.1 GI:23231144
JP2002509437-A/9.
unidentified
unidentified
unclassified.
1 (bases 1 to 1329)
Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F., Perron,H. and Mandrand,B.
Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses
Patent: JP 2002509437-A 9 26-MAR-2002;
BIO MERIEUX
OS Unidentified
FN JP 2002509437-A/9
PD 26-MAR-2002
PR 07-JUL-1998 JP 1999508255
PR 07-JUL-1997 PR 97/08816
PI GLAUCIA PARAHNOS BACCALA, FLORENCE KOMURIAN PRADEL, FREDERIC PI
PI MIREILLE SODOYER, CATHERINE OTT, FRANCOIS MALLET, HERVE PERRON,
PI BERNARD MANDRAND
PI C12N15/48, C12Q1/70, C07K14/15, A61K31/70
PC Strandedness: Single;
CC Topology: Linear;
CC Retroviral nucleic material and nucleotide fragments, in CC
CC associated with multiple sclerosis and/or
CC rheumatoid arthritis,
CC for
CC diagnostic, prophylactic and therapeutic uses FH Key
CC Location/Qualifiers
FT source 1. 1329
FT /organism='Unidentified'.
Location/Qualifiers

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	99.8	1329	6	BD136199 Retrovira
2	1326	99.8	1329	6	AR344389 Sequence
3	1326	99.8	1329	6	AX001030 Sequence
4	1316.4	99.1	163803	9	AC093531 Homo sapi
5	1076.8	81.0	10122	9	AY101590 Pongo pyg
6	1067	80.3	148711	9	AC133134 Homo sapi
7	1064.8	80.1	10124	9	AY101591 Pongo pyg
8	1036.8	78.0	2694	9	AF520478 Homo sapi
9	1036.8	78.0	2694	9	AF520490 Homo sapi
10	1036.8	78.0	2694	9	AF520506 Homo sapi
11	1036.8	77.9	2694	9	AF520530 Homo sapi
12	1035.2	77.9	2694	9	AF520484 Homo sapi
13	1035.2	77.9	2694	9	AF520488 Homo sapi
14	1035.2	77.9	2694	9	AF520494 Homo sapi
15	1035.2	77.9	2694	9	AF520496 Homo sapi
16	1035.2	77.9	2694	9	AF520498 Homo sapi
17	1035.2	77.9	2694	9	AF520504 Homo sapi
18	1035.2	77.9	2694	9	AF520510 Homo sapi
19	1035.2	77.9	2694	9	AF520514 Homo sapi

source		1. .1329		/organism="unidentified"		/mol_type="genomic DNA"		/db_xref="taxon:32644"	
Query Match		99.8%;		Score 1326;	DB 6;	Length 1329;			
Best Local Similarity		100.0%;		Pred. No. 0;	Mismatches	0;	Indels	0;	Gaps 0;
Matches 1329;		Conservative		0;					
ORIGIN									
QY	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTGTTTATTTTT	60						
DB	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTGTTTATTTTT	60						
QY	61	AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAGA	120						
DB	61	AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAGA	120						
QY	121	AATTTGAGATCGAATATATGTTAGAGCAGAGACCTTCAAAACACTGCACCTCGGGCCT	180						
DB	121	AATTTGAGATCGAATATATGTTAGAGCAGAGACCTTCAAAACACTGCACCTCGGGCCT	180						
QY	181	CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT	240						
DB	181	CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT	240						
QY	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTGTTAAGTTTGTCTCTCCAGAA	300						
DB	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTGTTAAGTTTGTCTCTCCAGAA	300						
QY	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA	360						
DB	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA	360						
QY	361	AATCTACCGTGGACCCCTGGACCGCTGCTAGACTATGCTCTGATGTTAATGACATTGA	420						
DB	361	AATCTACCGTGGACCCCTGGACCGCTGCTAGACTATGCTCTGATGTTAATGACATTGA	420						
QY	421	AGTACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCCAATTCAGTAGG	480						
DB	421	AGTACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCCAATTCAGTAGG	480						
QY	481	AAGCAGTTAGACAGTTGTACGCCAACCTCCCAACAGTACTTGGGTTTCTGTTTGA	540						
DB	481	AAGCAGTTAGACAGTTGTACGCCAACCTCCCAACAGTACTTGGGTTTCTGTTTGA	540						
QY	541	GGGTGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCNAGCCT	600						
DB	541	GGGTGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCNAGCCT	600						
QY	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTTGCAACTTAGCTCACACCG	660						
DB	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTTGCAACTTAGCTCACACCG	660						
QY	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720						
DB	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720						
QY	721	AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTGGGATATAAATCAGGCA	780						
DB	721	AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTGGGATATAAATCAGGCA	780						
QY	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTT	840						
DB	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTT	840						
QY	841	CACCTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTATGG	900						
DB	841	CACCTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTATGG	900						
QY	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGACCCGCT	960						
DB	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGACCCGCT	960						

QY	961	GCTGACTTTCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGTATCCAGTCCAGCGAGGT	1020						
DB	961	GCTGACTTTCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGTATCCAGTCCAGCGAGGT	1020						
QY	1021	ACCCATTGCCACTCCCAGTACAGGCTAAAGCTTGGCCATTTGCTCTGTGATCGCTAAGTGCC	1080						
DB	1021	ACCCATTGCCACTCCCAGTACAGGCTAAAGCTTGGCCATTTGCTCTGTGATCGCTAAGTGCC	1080						
QY	1081	TGGGTTGTCTCTAATAGAACTGAACTGTGTCTCACTGGGTTCCATGTTCTCTTCCATGAC	1140						
DB	1081	TGGGTTGTCTCTAATAGAACTGAACTGTGTCTCACTGGGTTCCATGTTCTCTTCCATGAC	1140						
QY	1141	CCAGGCTTCTAATAGAGCTATTAACACTCACCGCATGGCCCAAGATTTCCATTTCTTGGTA	1200						
DB	1141	CCAGGCTTCTAATAGAGCTATTAACACTCACCGCATGGCCCAAGATTTCCATTTCTTGGTA	1200						
QY	1201	TCTGTAGGCCAAGAACCCAGGTGACAGAAAGTGAGGCTTGCACCATTTGGGAAGTGG	1260						
DB	1201	TCTGTAGGCCAAGAACCCAGGTGACAGAAAGTGAGGCTTGCACCATTTGGGAAGTGG	1260						
QY	1261	CCCCTGTCATTTTGTAGTGGGCCACCACTCTCTGGAGCTGTGGGAGCAAGGATCCC	1320						
DB	1261	CCCCTGTCATTTTGTAGTGGGCCACCACTCTCTGGAGCTGTGGGAGCAAGGATCCC	1320						
QY	1321	CCAGTAACA 1329							
DB	1321	CCAGTAACA 1329							

RESULT 2

AR344389

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 1329;

Conservative

0;

Mismatches

0;

Indels

0;

Gaps

0;

1

TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTGTTTATTTTT

60

1

TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTGTTTATTTTT

60

61

AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAGA

120

61

AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAGA

120

121

AATTTGAGATCGAATATATGTTAGAGCAGAGACCTTCAAAACACTGCACCTCGGGCCT

180

121

AATTTGAGATCGAATATATGTTAGAGCAGAGACCTTCAAAACACTGCACCTCGGGCCT

180

181

CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT

240

181

CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT

240

241

TTTACTCTCTTTGGACCTGTATCTTCAACTTCTGTTAAGTTTGTCTCTCCAGAA

300

241

TTTACTCTCTTTGGACCTGTATCTTCAACTTCTGTTAAGTTTGTCTCTCCAGAA

300

301

TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA

360

301

TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA

360

361

AATCTACCGTGGACCCCTGGACCGCTGCTAGACTATGCTCTGATGTTAATGACATTGA

420

361

AATCTACCGTGGACCCCTGGACCGCTGCTAGACTATGCTCTGATGTTAATGACATTGA

420

421

AGTACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCCAATTCAGTAGG

480

421

AGTACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCCAATTCAGTAGG

480

481

AAGCAGTTAGACAGTTGTACGCCAACCTCCCAACAGTACTTGGGTTTCTGTTTGA

540

481

AAGCAGTTAGACAGTTGTACGCCAACCTCCCAACAGTACTTGGGTTTCTGTTTGA

540

541

GGGTGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCNAGCCT

600

541

GGGTGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCNAGCCT

600

601

ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTTGCAACTTAGCTCACACCG

660

601

ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTTGCAACTTAGCTCACACCG

660

661

ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC

720

661

ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC

720

721

AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTGGGATATAAATCAGGCA

780

721

AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTGGGATATAAATCAGGCA

780

781

TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTT

840

781

TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTT

840

841

CACCTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTATGG

900

841

CACCTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTATGG

900

901

CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGACCCGCT

960

901

CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGACCCGCT

960

Sequence 108 from patent US 6582703.

AR344389

AR344389.1

GI:33740330

Unknown.

Unknown.

Unclassified.

1 (bases 1 to 1329)

Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G., Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.

Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting

Patent: US 6582703-A 108 24-JUN-2003;

Location/Qualifiers

1. .1329

/organism="unknown"

/mol_type="mRNA"

Query Match

Best Local Similarity

Matches 1329;

Conservative

0;

Mismatches

0;

Indels

0;

Gaps

0;

1

TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTGTTTATTTTT

60

1

TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAAGAGGGGAACTGTTTATTTTT

60

61

AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAGA

120

61

AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAGA

120

121

AATTTGAGATCGAATATATGTTAGAGCAGAGACCTTCAAAACACTGCACCTCGGGCCT

180

121

AATTTGAGATCGAATATATGTTAGAGCAGAGACCTTCAAAACACTGCACCTCGGGCCT

180

181

CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT

240

181

CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT

240

241

TTTACTCTCTTTGGACCTGTATCTTCAACTGCTGTTAAGTTTGTCTCTCCAGAA

300

Db 241 TTTACTCTCTTGGACCTGTATCTTCAACTTCTTGTAAAGTTGTCTCTCCAGAT 300
301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTCTGATGTTAAATGACATTGA 420
361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTCTGATGTTAAATGACATTGA 420
421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG 480
421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG 480
481 AAGCAGTTAGAGAGTTGTCAGCAACCTCCCAACAGTACTGGTTTTCTGTTGAGA 540
481 AAGCAGTTAGAGAGTTGTCAGCAACCTCCCAACAGTACTGGTTTTCTGTTGAGA 540
541 GGTGGACTGAGAGCAGACTAGCTGGATTTCCTAGGCTGACTAAGATCCCNAGCCT 600
541 GGTGGACTGAGAGCAGACTAGCTGGATTTCCTAGGCTGACTAAGATCCCNAGCCT 600
601 ANCTGGGAAGGTGACCGCATCTCACTTTAAACATGGGGCTTGCAACTTAGCTCACACCGG 660
601 ANCTGGGAAGGTGACCGCATCTCTTAAACATGGGGCTTGCAACTTAGCTCACACCGG 660
661 ACCAATCAGAGAGCTCACTAAATGCTAATCAAGGCAAAACAGGAGGTAAAGCAATAGCC 720
661 ACCAATCAGAGAGCTCACTAAATGCTAATCAAGGCAAAACAGGAGGTAAAGCAATAGCC 720
721 AATCATCTATGCTGAGAGCAGCGGGAAGCAAGGATGGGATATAAACTCAGGCA 780
721 AATCATCTATGCTGAGAGCAGCGGGAAGCAAGGATGGGATATAAACTCAGGCA 780
781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTTTT 840
781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCCTCCCTCCCTTTT 840
841 CACTCTATTTCACTCTATTAATCATGCACTGCTCTTCTGGTCCGTTGTTTATGG 900
841 CACTCTATTTCACTCTATTAATCATGCACTGCTCTTCTGGTCCGTTGTTTATGG 900
901 CTCAAGCTGAGCTTTGTTGGCCATCCACACTGCTGTTGGCCACCGTCAACACCGCT 960
901 CTCAAGCTGAGCTTTGTTGGCCATCCACACTGCTGTTGGCCACCGTCAACACCGCT 960
961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAGGT 1020
961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAGGT 1020
1021 ACCCATGTCACCTCCCGATCAGGCTAAGGCTTGCCATGTTCTGATGGCTAAGTGC 1080
1021 ACCCATGTCACCTCCCGATCAGGCTAAGGCTTGCCATGTTCTGATGGCTAAGTGC 1080
1081 TGGGTTGTCCTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1081 TGGGTTGTCCTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1141 CCAAGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGTA 1200
1141 CCAAGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGTA 1200
1201 TCTGTAGGCAAGAACCCAGGTCAGAGAAAGTGGCTTGCCACCAATTTGGGAAGTGG 1260
1201 TCTGTAGGCAAGAACCCAGGTCAGAGAAAGTGGCTTGCCACCAATTTGGGAAGTGG 1260
1261 CCCACTGCAATTTGGTAGGCGCCCAACCATCTTTGGGAGCTGTTGGGCAAGGATCCC 1320
1261 CCCACTGCAATTTGGTAGGCGCCCAACCATCTTTGGGAGCTGTTGGGCAAGGATCCC 1320
1321 CCAAGTAACA 1329
|||||

Db 1321 CCAAGTAACA 1329
RESULT 3
AX001030
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match
Best Local Similarity 99.8%; Score 1326; DB 6; Length 1329;
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCCAAAAGAGGGGAACTGTTATTTT 60
Db 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCCAAAAGAGGGGAACTGTTATTTT 60
QY 61 AGGGGAAGAATCTGTGTAGTATGTTAACTGAATCTGAATCATTAAGAGTTAAAGA 120
Db 61 AGGGGAAGAATCTGTGTAGTATGTTAACTGAATCATTAAGAGTTAAAGA 120
QY 121 AATTTGAGATCCAAATATATGATAGAGCAGAGACCTTCAAAACACTGCACTGGGGCC 180
Db 121 AATTTGAGATCCAAATATATGATAGAGCAGAGACCTTCAAAACACTGCACTGGGGCC 180
QY 181 CTTAGCCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CTTAGCCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTTCTTAAAGTTGTCTCTTCCAGAT 300
Db 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTTCTTAAAGTTGTCTCTTCCAGAT 300
QY 301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
Db 301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
QY 361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTGCTGATGTTAAATGACATTGA 420
Db 361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTGCTGATGTTAAATGACATTGA 420
QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG 480
Db 421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG 480
QY 481 AAGCAGTTAGAGAGTTGTCAGCAACCTCCCAACAGTACTGGTTTTCTGTTGAGA 540
Db 481 AAGCAGTTAGAGAGTTGTCAGCAACCTCCCAACAGTACTGGTTTTCTGTTGAGA 540
QY 541 GGTGGACTGAGAGCAGACTAGCTGGATTTCCTAGGCTGACTAAGATCCCNAGCCT 600
Db 541 GGTGGACTGAGAGCAGACTAGCTGGATTTCCTAGGCTGACTAAGATCCCNAGCCT 600
QY 601 ANCTGGGAAGGTGACCGCATCTCACTTTAAACATGGGGCTTGCAACTTAGCTCACACCGG 660
|||||

Db	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	660
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	720
Db	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	720
Qy	721	AATCATCTATTGCTGAGAGACAGCGGAGGAGGACAGGATGGATATAAATCAGGCA	780
Db	721	AATCATCTATTGCTGAGAGACAGCGGAGGAGGACAGGATGGATATAAATCAGGCA	780
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTT	840
Db	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTT	840
Qy	841	CACCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTATGG	900
Db	841	CACCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTATGG	900
Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCT	960
Db	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCT	960
Qy	961	GCTGACTCCATCCCTTTGGATCCAGCAGAGGTGCTCACTGTGCTCCTGATCCAGCGAGGT	1020
Db	961	GCTGACTCCATCCCTTTGGATCCAGCAGAGGTGCTCACTGTGCTCCTGATCCAGCGAGGT	1020
Qy	1021	ACCCATTGCGCATCCCGATCAGGCTAAAGGCTTGGCATTTGCTGCAATGGCTAAGTGCC	1080
Db	1021	ACCCATTGCGCATCCCGATCAGGCTAAAGGCTTGGCATTTGCTGCAATGGCTAAGTGCC	1080
Qy	1081	TGGGTTTGCTTAATAGACTGAACACTGCTCCTGGGTCCATGTTCTTCCATGAC	1140
Db	1081	TGGGTTTGCTTAATAGACTGAACACTGCTCCTGGGTCCATGTTCTTCCATGAC	1140
Qy	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGTA	1200
Db	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGTA	1200
Qy	1201	TCTGTGAGCCGAAGAACCCAGGTCAGAGAAGTGAAGCTTGCACCATTTGGGAAGTGG	1260
Db	1201	TCTGTGAGCCGAAGAACCCAGGTCAGAGAAGTGAAGCTTGCACCATTTGGGAAGTGG	1260
Qy	1261	CCCACTGCCATTTTGTAGCGGCCACCACTCTTGGAGCTGTGGGAGCAAGATCCC	1320
Db	1261	CCCACTGCCATTTTGTAGCGGCCACCACTCTTGGAGCTGTGGGAGCAAGATCCC	1320
Qy	1321	CCAGTAACA 1329	
Db	1321	CCAGTAACA 1329	
RESULT 4			
AC093531/c			
LOCUS			
DEFINITION			
AC093531			
ACCESSION			
AC093531.2			
GI:16945981			
HTG.			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 163803)			
DOE Joint Genome Institute and Stanford Human Genome Center.			
AUTHORS			
TITLE			
Direct Submission			
JOURNAL			
Unpublished			
REFERENCE			
2 (bases 1 to 163803)			
DOE Joint Genome Institute.			
AUTHORS			
TITLE			
Direct Submission			
JOURNAL			
Submitted (31-AUG-2001) Production Sequencing Facility, DOE Joint			
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
3 (bases 1 to 163803)			
REFERENCE			
DOE Joint Genome Institute and Stanford Human Genome Center.			

TITLE		Direct Submission	
JOURNAL		Submitted (16-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell	
COMMENT		Drive, Walnut Creek, CA 94598, USA	
		On Nov 16, 2001 this sequence version replaced gi:15383820.	
		Draft Sequence Produced by DOE Joint Genome Institute	
		www.jgi.doe.gov	
		Finishing Completed at Stanford Human Genome Center	
		www.shgc.stanford.edu	
		Quality: Phrap Quality >=40 99.8% of Sequence;	
		Estimated Total Number of Errors is 0.3.	
FEATURES		Location/Qualifiers	
source		1..163803	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/chromosome="5"	
		/clone="RP11-405L7"	
ORIGIN		Query Match 99.1%; Score 1316.4; DB 9; Length 163803;	
		Best Local Similarity 99.3%; Pred. No. 0;	
		Matches 1320; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGCTTAACCGCCAAAGAGGGGAACTGTTTATTTT	60
Db	117604	TTAAAATTGAAGAGCTTTAGACTTGCTTAACCGCCAAAGAGGGGAACTGTTTATTTT	117545
Qy	61	AGGGGAAGATGCTGTTAGTAGTATGTTAATCAATCTGGAATCATTAATCAGAAAGTTAAAGA	120
Db	117544	AGGGGAAGATGCTGTTAGTAGTATGTTAATCAATCTGGAATCATTAATCAGAAAGTTAAAGA	117485
Qy	121	AATTTGAGATCGAATATAATGTTAGAGCAGAGGACCTTTCAAACACATGACCCCTGGGCT	180
Db	117484	AATTTGAGATCGAATATAATGTTAGAGCAGAGGACCTTTCAAACACATGACCCCTGGGCT	117425
Qy	181	CCTCAGCAATGGAATGCCCTGGACTCTCCCTCTTAGGACCTTAGCAGCTATAATTT	240
Db	117424	CCTCAGCAATGGAATGCCCTGGACTCTCCCTCTTAGGACCTTAGCAGCTATAATTT	117365
Qy	241	TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTCCAGAA	300
Db	117364	TTTACTCTCTTTGGACCCCTGTATCTTAACTCTCTTGTAAAGTTTGTCTCTCCAGAA	117305
Qy	301	TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCATGACTAA	360
Db	117304	TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCATGACTAA	117245
Qy	361	AATCTACCGTGACCCCTGGACCCGCTGCTAGACTATGCTCTGATGTTAATGACATTGA	420
Db	117244	AATCTACCGTGACCCCTGGACCCGCTGCTAGACTATGCTCTGATGTTAATGACATTGA	117185
Qy	421	AGTCACCCCTCCGAGGAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGG	480
Db	117184	AGTCACCCCTCCGAGGAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGG	117125
Qy	481	AAGCAGTTAGACGAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA	540
Db	117124	AAGCAGTTAGACGAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA	117065
Qy	541	GGGTGGACTGAGACAGGACTAGCTGATTTCTTAGGCTGACCTAAGAAATCCCAAGCT	600
Db	117064	GGGTGGACTGAGACAGGACTAGCTGATTTCTTAGGCTGACCTAAGAAATCCCTAAGCT	117005
Qy	601	ANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	660
Db	117004	AGCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	116945
Qy	661	ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	720
Db	116944	ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	116885
Qy	721	AATCATCTATTGCTGAGAGACAGCGGAGGACAGGATGGGATATAAATCAGGCA	780

Db 9335 TGGGGACTGAGAGCAGACTAGCTGATTTCTTAGCCCTACTAAGAAATCCCTAAGCCT 9394
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGGCAACTTAGCTCACACCG 660
Db 9395 AGCTGGGAAGGTGACCAATCCATCTTTAAACACGGGGCTTGGCAACTTAGCTCACACCTG 9454
Qy 661 ACCAATCAGAGAGCTCATAAATCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC 720
Db 9455 ACCAATCAGAGAGCTCATAAATCTAATTAGGCAAAAACAGGAGGTAAAGCAATAGCC 9514
Qy 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAATCAGGCA 780
Db 9515 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAATCAGGCA 9574
Qy 781 TTCAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGATGAGGAGCTGTGTTT 840
Db 9575 TTCGAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGATGAGGAGCTGTGTTT 9634
Qy 841 CACTCTATTTCATCTATTAAATCATCAACAGCACTCTTCTGGTCCGCTGTTTATGG 900
Db 9635 CATGCTATTTCATCTATTAAATCTTGAACATGCACTCTTCTGGTCCGCTGTTTATGG 9694
Qy 901 CTCAGCTGAGCTTTTGTTCGCAATCCCACTGCTGTTTGGCCACCGTCACAGACCCGCT 960
Db 9695 CTTGAGCTGAGCTTTTGTCTGCTGTCACCACTGCTGTTTGGCCACCGTCACAGACCCGCT 9754
Qy 961 GCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGGAGGT 1020
Db 9755 ACTGACTCCCATCCCTCCGGATCTTGCAGGGTGTCCACTGTGTTCTGATCCAGGAGGT 9814
Qy 1021 ACCCAATGCCACTCCCGATCAGGCTAAAGGCTTGCATTGTTCTTCGATGCTTAAGTGC 1080
Db 9815 GCCCAATGCGCTCCCAATCAGGCTAAAGGCTTGCATTGTTCTTCGATCATAAGTGC 9874
Qy 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCATGGTTCCTTCCATGAC 1140
Db 9875 CAGGTTGCTCTTAATGAGCTGAACACTAGTCACTGGGTTCACAGGTTCTCTTCTGTGAC 9934
Qy 1141 CCACGGCTTCTAATAGAGTATAACACTACCGCATGCGCCAGAGTTCATTCCTTGTA 1200
Db 9935 CCACAGCTTCTAATAGAGTATAACACTACCACTACCACTGCGCCAGGTTCCATTCCTT-GAA 9993
Qy 1201 TCTGTGAGCCAAAGAACCCAGCTCAGAGAAGTGGAGCTTGCCACCAATTTGGGAAGTGG 1260
Db 9994 TCATGAGGCCAAGAACCCAGCTCAGAGAATGAGGCTTGCCACCACTTTGGAGCAG 10053
Qy 1261 CCCACTGCCATTTGGTAGCGGCCCAACCACTTTGGAGCTGTGGAGCAAGATCCC 1320
Db 10054 CCGGCCACCATCTTGAAGCAGCTCACCACTCTTGGAGCTGTGTGAGCAAGGACCCC 10113
Qy 1321 CCAGTAACA 1329
Db 10114 CAAGTAACA 10122

RESULT 6
AC133134
LOCUS AC133134 148711 bp DNA linear PRI 28-NOV-2002
DEFINITION Homo sapiens chromosome 5 clone RP11-120B7, complete sequence.
ACCESSION AC133134
VERSION AC133134.2 GI:25815299
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 148711)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 148711)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission
JOURNAL Submitted (07-SEP-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 148711)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (28-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Nov 28, 2002 this sequence version replaced gi:22758291.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.7.
FEATURES
source
1. 148711
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-120B7"
ORIGIN
Query Match 80.3%; Score 1067; DB 9; Length 148711;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 1175; Conservative 0; Mismatches 143; Indels 4; Gaps 2;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTGTTATTTT 60
Db 99272 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTGTTATTTT 99331
Qy 61 AGGGAAGAAATGCTTTAGTATGTTAATCAATCTGGAATCATTAATGAGAAGTTAAAGA 120
Db 99332 AAAGGAAGATGCTTTATGTTATGTTAATCAATCAGAAATCGTCTGAGAAGTTAAAGA 99391
Qy 121 AATTGAGATCGAATATAATGTAGAGCAGAGAGCTTCAAAACACTGACCCCTGGGGCT 180
Db 99392 AATTCGAAATCAAAATCAACCGTAGAGCAAAAGAGCTTCAAAACGTTGGGACACTGAGGCT 99451
Qy 181 CCTCAGCAATGGATGCCCTGGAGCTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 99452 CCTCAGCAATGGATGCCCTGGATTTTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 99511
Qy 241 TTTACTCTCTTTTGAGACCTGTATCTTCAACTTCTCTGTTAAGTTGTCTCTTCAGAAAT 300
Db 99512 GCTACTCATTTTAGACCTCTATCTTTAACTCTCTTAAAGTTGTCTCTTCAGAAAT 99571
Qy 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATGCCAGTCCATGACTAA 360
Db 99572 CGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATGCCAGTCCATGACTAA 99631
Qy 361 AATCTACCGTGAGCCCTGGAGCCGCTGTAGACTATGCTGTGATGTTAATGACATTGA 420
Db 99632 GATCTACCGCGGACCCCTGGAGCTGGCTGTGCTGTAGCCCATGCTGTGATGTTAATGACATTGA 99691
Qy 421 AGTCACCCCTCCCGAGGAAATCTCAACTGSCAAACCCCTACTACACTCCAATTCAGTAGG 480
Db 99692 AGGCACCCCTCCAGAGGAAATTTCAACTGSCAAACCCCTACTACACCCTAATTCAGCAGG 99751
Qy 481 AAGCAGTTAGAGCAGTGTGTAGCCCAACCTCCCCAACAGTACTTGGGTTTCCCTGTTGAGA 540
Db 99752 AAGCAGTTAGAGTGTGTGTAGCCCAACCTCCCCAACAGTACTTGGGTTTCCCTGTTGAGA 99811
Qy 541 GGGTGGACTGAGAGCAGGACTAGTGTGATTTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 99812 GGGGTACTGAGAGCAGGACTAGTGTGATTTTCTTAGGCCGACTAAGAAATCCCTAAGCCT 99871
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGGCAACTTAGCTCACACCGG 660
Db 99872 AGCTGGGAAGGTGACCGCATCCACTTTTAAACATGGGCTTGGCAACTTAGCTCACACCGG 99931
Qy 661 ACCAATCAGAGAGCTCATAAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC 720

[illegible]

301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGTCAGTCCATGACTAA 360
 1723 CGAAGCTGTAAACTA-----CAATGAGCCCAAGATGTCAGTCCAGACTAA 1770
 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGGA 420
 1771 GATCTACCGACAGCCCTGGACCGGCTGCTAGCCGACGATCTGATGTTAATGACATCAA 1830
 421 AGTCACCCCTCCGAGGAATCTCACTGACAAACCCCTACTACACTCCAATTCAGTAGG 480
 1831 AGGCACCCCTCTGAGGAATCTCACTGACAAACCTCTACTACGCCCAATTCAGCAGG 1890
 481 AAGCAGTTAGACAGTTGTGACCAACCTCCCAACAGTACTTGGTTTCTCTGTTGAGA 540
 1891 AAGCAGTTAGACGGTGTGCGCCCAACCTCCCAACAGCAGTACTGTTTCTCTGTTGAGA 1950
 541 GGGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGATGAAGATCCCAAGCCT 600
 1951 TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGATGAAGATCCCAAGCCT 2010
 601 ANCTGGGAAGTACCGGATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCG 660
 2011 AGCTGGGAAGTACCGGATCCATCCACCTTTAAACACGGGGCTTGCACCTTAGCTCACACCTG 2070
 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC 720
 2071 ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAAACAGGAGGTAAGCAATAGCC 2130
 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
 2131 AATCATCTATTGCTGAGAGCAGCGGAGGCAATGATCGGGATATAAAACCCCAAGTC 2190
 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCTCCCTCAATGATGAGGAGCTCTGTTT 840
 2191 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCTCCCTCAATGATGAGGAGCTCTGTTT 2250
 841 CACTCTATTCTACTTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTATGG 900
 2251 CATGCTATTCTACTTATTAAATCTTGCAACTGCACTCTTCTGGTCCGTTTATGG 2310
 901 CTCAAGCTGAGCTTTTGTTCGCATCCACCTGCTGTTTGGCCACCGTCACAGACCCGT 960
 2311 CTGAGCTGAGCTTTTGTTCGCATCCACCTGCTGTTTGGCCACCGTCACAGACCCGT 2370
 961 GCTGACTTCATCCCTTTGGATCCAGCAGAGTGCACCTGCTCTCTGATCCAGCAGGT 1020
 2371 GCTGACTTCATCCCTTTGGATCCAGCAGAGTGCACCTGCTCTCTGATCCAGCAGGT 2430
 1021 ACCATTGGCACTCCCGATCAGGCTAAAGCTTGGCATTGTTCTGCTGCTTAAGTGC 1080
 2431 ACCATTGGCACTCCCGATCAGGCTAAAGCTTGGCATTGTTCTGCTGCTTAAGTGC 2490
 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGTTCTTCTCCATGAC 1140
 2491 TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGTTCTTCTCTGAC 2550
 1141 CACGCTTCTTAATAGAACTGAACACTGGGTTCCATGTTCTTCTCTGTTGTTA 1200
 2551 CACGCTTCTTAATAGAACTGAACACTGGGTTCCATGTTCTTCTCTGTTGTTA 2609
 1201 TCTGTGAGCCCAAGAACCCCGATCAGAGAAAGTGGGCTTGGCACCATTGGGAAG 1257
 2610 TCCATAAGCCCAAGAACCCCGATCAGAGAAAGTGGGCTTGGCACCATTGGGAAG 2666

RESULT 9
 AF520489S2 2694 bp DNA linear PRI 11-FEB-2004
 LOCUS Homo sapiens individual 22 allele A, envelope glycoprotein gene,
 DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
 AF520490
 ACCESSION
 VERSION
 AF520490.1 GI:33410948
 KEYWORDS

2 of 2
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
 Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
 TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
 in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBMED 14757826
 REFERENCE 2 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O. and Oriol,G.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée
 d'Italie, Lyon 69364 cedex 07, France
 FEATURES
 Location/Qualifiers
 1..2694
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="individual 22 allele A"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-q22"
 /sex="male"
 /cell_type="PBMC's"
 /note="Caucasian"
 /endogenous_virus="HERV-W"
 43..57
 /note="splice acceptor site"
 284..1900
 /note="syncytin"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="AAQ17567.1"
 /db_xref="GI:33410949"
 /translation="MALPYHFLFTVLLPSFTLTAPPCCRCMTSSSPYOEFLLRMORP
 GNIDAPSVRLSKGTPPTAHTMPNCHYHSATLQMHANTHYMTGKMLNPSPCGGLGV
 TVCTYFTGMSDGGVQDQAKREKVEVLSQLTRVHGTSPPYKGLDLSKLHEILRT
 HTRLVSLFNTTLGLHVEASQNPNCICLPLNFRPYVSIIPVEQNNPSTSEINTSV
 LVGLVLSLEIHTSNLTCVKPSTNTYTNQYVIRKPRKRVPLPFTVIGAGVGLGALG
 YRLNGSSSEFLSPFVPMPTIYEQDLNYSIKRPNKRVPLPFTVIGAGVGLGALG
 TGIGITSTOFYKLSQELNMDVADSLVTLQDLNLSAAVVLQNRALDLDLTAEE
 RGCTCLFGECCYVNGSGIVTEKVEIRDIORAEELRTGPKGLLSQWMPVTLTP
 FLGPAALIIILLPGLFIFNLLVNFVSSRIENAVKLQMEPKMQSKTIYRRLDRPAS
 RSDVNDIKGTPPEISAAQPLLRPNASGSS"
 1705..1720
 /note="splice acceptor site"
 1959..2694

Query Match 78.0%; Score 1036.8; DB 9; Length 2694;
 Best Local Similarity 90.2%; Pred. No. 9.7e-305;
 Matches 1134; Conservative 0; Mismatches 110; Indels 13; Gaps 2;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCTGAAAGAGGGGAACTGTTATTTT 60
 DB 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCTGAAAGAGGGGAACTGTTATTTT 1482
 QY 61 AGCGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAAGTTAAAGA 120
 DB 1483 AGCGGAAGAAATGCTGTTAGTATGTTAATCAATCCGGAATCGTCACTGAGAAAAGTTAAAGA 1542
 QY 121 AATTTGAGATCGAATATATATGATAGCAGCAGAGACCTTCAAAACACTGCACCTGGGGCCT 180
 DB 1543 AATTCGAGATCGAATATACACACGATAGCAGCAGAGAGCTTCGAAACACTGGACCTGGGGCCT 1602
 QY 181 CCTGACCAATGATGATCCCTGACCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT 240
 DB 1603 CCTGACCAATGATGATGATCCCTGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT 1662

complete cds, and 3' long terminal repeat, complete sequence.	
AF520506	AF520506.1 GI:33410980
VERSION	
KEYWORDS	
SOURCE	
SEGMENT	
ORGANISM	
2 of 2	
Homo sapiens (human)	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 2694)	
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.	
The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology	
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
14757826	
2 (bases 1 to 2694)	
Mallet, F., Bouton, O. and Oriol, G.	
Direct Submission	
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France	
Location/Qualifiers	
1. .2694	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/isolate="individual 37 allele A"	
/db_xref="taxon:9606"	
/chromosome="7"	
/map="7q21-q22"	
/sex="male"	
/cell_type="PBMC's"	
/notes="endogenous_virus: HERV-W"	
43..57	
/note="splice acceptor site"	
284..1900	
/note="syncytin"	
/codon_start=1	
/product="envelope glycoprotein"	
/protein_id="AA017575.1"	
/db_xref="GI:33410981"	
/translation="MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPVQSFRLMRQRP GNDAPSYRLSKGTPTFTAHMPKNCYHSAILCMHANTHYTGKMINSPCGGLGV TCVWYPTQSGMDGGVQDOAREKHVKVISQLTRVHGSHSTYKGLDLSKLHETLRT HTLVLNFTLTGLHEVSAQNTNCWICLPLNFRPYVSPVPEQWNNFSTEINTTV LVLPLVNLBIHTSNLTQVKFNSTVITNSQICIRWYPTQIVCLPSGIFPVGCTSA YRCLNGSSSMCLFLVPPMTIYTEDLYNVISKPNKRVPLIPVIGAGVLGALG TGGIITTSQFYKLSOELNGDMERVADSLVTLQDOLSLAAVLQNRALDILLTAE RGGTCLFLGSECCYYVNSQGIYVEKVEIRDRIQORAEELRNTGPGWLLSQWMPWILP FLGPLAAIIILLFLGPCIIFNLLVNFSSRIEAVKLQMEPKMQSKTKIYRRPLDRPAS RSDVNDIKGTPTPEEISAAQPLLRPNASGS"	
1705..1720	
/note="splice acceptor site"	
1959..2694	
misc_feature	
LTR	
ORIGIN	
Query Match 78.0%; Score 1036.8; DB 9; Length 2694;	
Best Local Similarity 90.2%; Pred. No. 9.7e-305;	
Matches 1134; Conservative 0; Mismatches 110; Indels 13; Gaps 2;	
Qy	1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGGAACCTGTTTATTTT 60
Db	1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAGAGGGGGAACCTGTTTATTTT 1482
Qy	61 AGGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCATCTAGAGAAGTTAAAGA 120
Db	1483 AGGGGAAGAATGCTGTATTTATGTTAATCAATCCGAATCGTCACTGAGAAGTTAAAGA 1542
Qy	121 AATTGAGATCGAATATTAATGTAGACGAGAGCCTTCAAAACACTGCACCTGGGCCT 180
Db	1543 AATTGAGATCGAATATCAACGCTAGACGAGAGAGCTTCGAAACACTGCACCTGGGCCT 1602
Qy	181 CCTCAGCAATGATGCTCCCTCTCTCCCTCTCTTAGGACCTCTCTAGCAGCTATAATATT 240

1603 Db CTCAGCCATGATGCTGCTGATCTCTCCCTCTTAGGACTCTAGCAGCTATATATTT 1662
 241 QY TTTACTCTCTTTGGACCTCTATCTTCAACTTCTCTGTTAAAGTTGCTCTTCCAGAAAT 300
 1663 Db GCTACTCTCTTTGGACCTGATCTTTAACTCTCTTAACTTGTCTCTTCCAGAAAT 1722
 301 QY TGAAGCTGTAAGCTTACAATAGTTCTTCAAAATGGAACCCAGATGCCAGTCCATCACTAA 360
 1723 Db CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATGCCAGTCCCAAGACTAA 1770
 361 QY AATCTACCTGTGACCCCTCGACCGGCTCTAGACTATCTCTGATGTTAAATGACATTTGA 420
 1771 Db GATCTACCCGACACCCCTGGACCGGCTCTAGCCACGATCTGATTAATGACATCAA 1830
 421 QY AGTCACCCCTCCGAGGAATCTCAACTGCAACAAACCCCTACTACACTCCAAATCTAGTAGG 480
 1831 Db AGGCACCCCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
 481 QY AAGCAGTTAGAGAGTTGTGACGCAACCTCTCCCAACAGTACTTGGGTTTCTCTGTGAGA 540
 1891 Db AAGCAGTTAGAGAGTTGTGCGCAACCTCTCCCAACAGTACTTGGGTTTCTCTGTGAGA 1950
 541 QY GGTGAGCTGAGACAGAGCTAGCTGGAATTTCTAGGCTGACTAAGAAATCCCAAGCCT 600
 1951 Db TGGGGAGCTGAGACAGAGCTAGCTGGAATTTCTAGGCTGACTAAGAAATCCCAAGCCT 2010
 601 QY ANCTGGGAAGGTGACCGCATCTCTTTAAACATGCGGCTTGCAACTTAGCTCACACCGG 660
 2011 Db AGCTGGGAAGGTGACCACTCCACCTTTAAACAGCGGCTTGCAACTTAGCTCACACCTG 2070
 661 QY ACCAATCAGAGAGCTCACTAAATGCTATTCAGGCAAAACAGGAGTTAAGCAATAGCC 720
 2071 Db ACCAATCAGAGAGCTCACTAAATGCTATTCAGGCAAAACAGGAGTTAAGCAATAGCC 2130
 721 QY AATCATCTATTGCTCAGAGCAGCGGGAAGCAAGGATTCGGGATATAAACTCAGGCA 780
 2131 Db AATCATCTATTGCTCAGAGCAGCGGGAAGCAAGGATTCGGGATATAAACTCAGGCA 2190
 781 QY TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATGTTATGGAGCTCTCTTTT 840
 2191 Db TTGAGCCGCAACAGCAACCCCTTTGGGTCCCTCCCATGTTATGGAGCTCTCTTTT 2250
 841 QY CACTCTATTCTCTATTAAATCATGCACTGCACTCTCTGCTCGGTCTTTTATGG 900
 2251 Db CATGCTATTCTCTATTAAATCTTGCACTGCACTCTCTGCTCGGTCTTTTATGG 960
 901 QY CTTAAGCTGAGCTTTTGTTCGCCATCCACCTGCTGTTTGGCCACGTCACAGACCGCT 960
 2311 Db CTTGAGCTGAGCTTTTGTTCGCCATCCACCTGCTGTTTGGCCACGTCACAGACCGCT 2370
 961 QY GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGGT 1020
 2371 Db GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGGT 2430
 1021 QY ACCATTTGCACTCCCGATCAGGCTTAAAGCTTGCATGTTTCTCTGATGCTTAAAGTGC 1080
 2431 Db ACCATTTGCACTCCCGATCAGGCTTAAAGCTTGCATGTTTCTCTGATGCTTAAAGTGC 2490
 1081 QY TGGGTTTGTCTTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 2491 Db TGGGTTTGTCTTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2550
 1141 QY CCACGGCTTCTAATAGAGCTATAAAGCTACCGCATGCGCCCAAGATTCCTATCTTTGGTA 1200
 2551 Db CCACGGCTTCTAATAGAGCTATAAAGCTACCGCATGCGCCCAAGATTCCTATCTTTGGTA 2609
 1201 QY TCTGTGAGGCAAGAACCCAGCTCAGAGAAAGTGGGCTTGGCACCATTGGGAG 1257
 2610 Db TCCATAAGGCAAGAACCCAGCTCAGAGAAAGTGGGCTTGGCACCATTGGGAG 2666

RESULT 11

AF520529S2
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM

AF520529S2
 Homo sapiens individual 78 allele A, envelope glycoprotein gene,
 complete cds, and 3' long terminal repeat, complete sequence.
 AF520530
 AF520530.1 GI:33411028
 2 of 2
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS

1 (bases 1 to 2694)
 Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
 Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
 The endogenous retroviral locus ERVW1 is a bona fide gene involved
 in hominoid placental physiology
 Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 14757826

REFERENCE
 AUTHORS

2 (bases 1 to 2694)
 Mallet,F., Bouton,O. and Oriol,G.
 Direct Submission
 Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
 d'Italie, Lyon 69364 cedex 07, France

FEATURES
 source

1.2694
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="individual 78 allele A"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-q22"
 /sex="female"
 /cell_type="PBMC's"
 /note="Caucasian"
 endogenous_virus: HERV-W"
 43..57
 /note="splice acceptor site"
 284..1900
 /note="syncytin"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="AAQ17587.1"
 /db_xref="GI:33411029"
 /translation="MALPYHIFLTVLLPSFTLTAPPCCMTSSPYQSEFLWRMORP
 GNIDAPYSRSLSKGPTPTTAHTMPNRCNTHSATLCMHANTHWTKMINPSCPGGLGV
 TVCTYFTQTMGSDGGVQDQAREKHVEVISQLTRVHGTSSPYKGLDLSKHEITRTV
 HTVLSLNTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPIVPEQNNFSTEINTTSV
 LVGLVSLNLEIHTHTSLNLTCKFSNTYTTNSQICIRWVTPPTQICLPSGIFVCGTSA
 YCLNGSSSCFLSLVPPMTIYTEQDLYNVYISKPRNRPVILPFVIGAGVLGALG
 TGIGITTSQFYKLSQELNGDMERVADSLTLODOLNSLAUVLONRRALDOLLAE
 RGCTCLFLGEECCYVNSGIVTEKVEIRDRIRKAEELRNTGPGWLLSQWMPHILP
 FLGPLAAIILLFLPGFCIFNLNLFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
 RSDVNDIKGTPEEISAAQPLLRPSAGSS"
 1705..1720
 /note="splice acceptor site"
 1959..2694

misc_feature

CDS

misc_feature

LTR

ORIGIN

Query Match 78.0%; Score 1036.8; DB 9; Length 2694;
 Best Local Similarity 90.2%; Pred. No. 9,7e-305;
 Matches 1134; Conservative 0; Mismatches 110; Indels 13; Gaps 2;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTTATTTT 60
 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGAACTGTTATTTT 1482

QY 61 AGGGGAAGATGCTGTTAGTATGTTTAACTCAATCTGAATCATTACTGAGAAAGTAAAGA 120
 Db 1483 AGGGGAAGATGCTGTTATGTTTAACTCAATCCGGAATCGTCACTGAGNAAGTTAAGA 1542

QY 121 AATTTGAGATCGAATAATAATAGTAGACGAGGAGACCTTCAAAACACTGCACCTGGGCGCT 180


```
Db 1423 TCATAATCGAAGAGCTTTAGACTTGTCTAAACGCTGAAGAGGGGAAACCTGTTTATTTT 1482
Qy 61 AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCATATTCTGAGAAGTTAAAGA 120
Db 1483 AGGGGAAGAATGCTGTTATATTATGTTAATCAATCTGGAATCTGCTACCTGAGAAGTTAAAGA 1542
Qy 121 AATTGAGATCGATATATCTAGAGCAGAGAGCCTTCARAAACACTGACACCTCGGGCCT 180
Db 1543 AATTGAGATCGAATACAACTGAGAGCAGAGAGCTTCGAARCACTGACACCTCGGGCCT 1602
Qy 181 CCTCAGCCAATGGATGCGCTGACTCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 1603 CCTCAGCCAATGGATGCGCTGACTCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 1662
Qy 241 TTTACTCTCTTTGAGCCCTGATCTTCAACTCTTCTGTTAAGTTTGTCTCTCCAGNAT 300
Db 1663 GCTACTCTCTTTGAGCCCTGATCTTAAACCTCTCTGTTAACTTGTCTCTCCAGNAT 1722
Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCACTGCATGACTAA 360
Db 1723 CGAAGCTGTAAGCTA-----CAATGGAGCCCAAGATGCACTGCATGCACTAA 1770
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCA 1830
Qy 421 AGTCACCCCTCCGAGGAATCTCACTGACACACCCCTACTACTCACTCAATTCAGTAGG 480
Db 1831 AGGCACCCCTCTCGAGGAATCTCACTGCTGCAACCTCTACTACGCCCAATTCAGCAGG 1890
Qy 481 AAGCAGTTTAGACAGCTGTGTCAGCAACCTCCCAACAGTACTTGGTGTCTCTGTTGAGA 540
Db 1891 AAGCAGTTTAGACCGGTGCTGCGCCAACTCCCAACAGACTTAGTGTCTGTTGAGA 1950
Qy 541 GGGTGGACTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTGAATCCCNAGCCT 600
Db 1951 TGGGGGACTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTGAATCCCNAGCCT 2010
Qy 601 ANCTGGGAAGTGACCGCATCTTTAAACATGGGCTTGCACTTAGCTACACCCG 660
Db 2011 AGCTGGGAAGTGACCATCATCCCTTTAAACACGGGGCTTGCACTTAGCTACACCTG 2070
Qy 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAAAGGAGGTAAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAAAGGAGGTAAAGCAATAGCC 2130
Qy 721 AATCATCTATTCCCTGAGAGCAGCGGGAAGCAAGGATGGGATATAAATCAGGCA 780
Db 2131 AATCATCTATTGCTGAGAGCAGCAGGAGGGAATGATCGGGATATAAATCAGGCA 2190
Qy 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCATGTTATGGAGCTCTGTTT 840
Db 2191 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTT 2250
Qy 841 CACTCTATTCTACTTATTAAATCATGCAACTGCACTCTCTTGGTCCGTTGTTTATGG 900
Db 2251 CATGCTATTCTACTTATTAAATCTTGCACTGCACTCTCTGTTCCATGTTTCTACGG 2310
Qy 901 CTCAGCTGAGCTTTTGTTCGCAATCCACCACTGCTGTTTGGCCAGCGTCAAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTCGCTCGCATCCACCACTGCTGTTTGGCCAGCGTCAAGACCCGCT 2370
Qy 961 GCTGACTCCATCCCTTTGGATCCAGCAGAGGTCCACTGTGCTCTCTGATCCAGCGAGT 1020
Db 2371 GCTGACTCCATCCCTCTGGATCATGAGGGTGTCCGCTGTGCTCTCTGATCCAGCGAGC 2430
Qy 1021 ACCCATTTGCCATCCCGATCAGGCTTAAAGGCTTGCCATTTGTTCTGCTAGTGCCTAAAGTCC 1080
Db 2431 ACCCATTTGCCCTCCCAATTTGGCTTAAAGGCTTGCCATTTGTTCTGCTAGTGCCTAAAGTCC 2490
Qy 1081 TGGGTTTGTCTTAATAGAACTGAACCTGCTCAGTGGGTCCATGGTCTCTCTCCATGAC 1140
Db 2491 TGGGTTTGTCTTAATAGAACTGAACCTAGTCACTGGGTCCATGGTCTCTCTCTGTTGAC 2550
```

```
Qy 1141 CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGTA 1200
Db 2551 CCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTT-GAA 2609
Qy 1201 TCTGTGAGGCCAAGAACCCCGAGGTGAGAAAGTGAGGCTTGCCACCATTTGGGAAG 1257
Db 2610 TCCATAGGCCAAGAACCCCGAGGTGAGAAAGTACAGAGGCTTGCCACCATCTTGGGAG 2666

RESULT 14
AF520493S2 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 24 allele A, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520494
VERSION AF520494.1 GI:33410956
KEYWORDS
SEGMENTS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
JOURNAL 14757826
PUBMED 2 (bases 1 to 2694)
REFERENCE Mallet,F., Bouton,O. and Oriol,G.
AUTHORS Direct Submission
TITLE Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
JOURNAL CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 24 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="Asian
endogenous_virus: HERV-W"
misc_feature 43..57
/note="splice acceptor site"
CDS 284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17569.1"
/db_xref="GI:33410957"
/translation="MALPYHIFLFTVLLPSFTLAPPRCMCTSSSPQBFWMQRP
GNIDAPYRSLSKGPTPTFAHTHMPNRCYHSATLCMHANTHYTKMINDSPCGGLGV
TVCKYFTQTMDSGGGVQDAAREKHVKVISOLTRVHGTSPPYKGLDLSKLHETLR
HTRLVSNLEITHSTNLTGCVKFNPTNCHICLPNFRPVYSIPVPSMNNFSTEINTSV
LVGLVSNLEITHSTNLTGCVKFNPTNCHICLPNFRPVYSIPVPSMNNFSTEINTSV
YRCINGSESGKFLSVLPVPMITVTEQDLYSVISKPNKRPVILPFIAGVLCALG
TLCGGITTSQFYKYSQELQNDMERVADSLVTLDQNLNSLAAVVLQNRALDLTAE
RGCTPLGECYVYKYSQELQNDMERVADSLVTLDQNLNSLAAVVLQNRALDLTAE
FGPLAAIILLLFGPCIFNLLVNFVSRIEAVKLQMEPKMQSKTKIYRRPLDRPAS
RSDYNDIKGTPPEISAAQPLLRPSAGSS"
misc_feature 1705..1720
/note="splice acceptor site"
LTR 1959..2694

Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
```

Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;

QY 1 TCAAAATCGAAGAGCTTTAGACTGCTAAACCCGCAAAAGAGGGGAACTGTTTATTTT 60
 Db 1423 TCAAAATCGAAGAGCTTTAGACTGCTAAACCCGCTGAAAGAGGGGAACTGTTTATTTT 1482

QY 61 AGGGGAGAGTCTGTTAGTATGTTAAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
 Db 1483 AGGGGAGAGTCTGTTATTTATTTAATCAATCCGAAATCTGCTAGAGAAAGTTAAAGA 1542

QY 121 AATTTGAGATCGAATATATATAGAGCAGAGGACCTTCAAAACACTGCAACCTCGGGCCT 180
 Db 1543 AATTCGAGATCGAATACACAGTAGAGCAGAGGAGCTTCGAAACACTGGACCCCTGGGGCT 1602

QY 181 CCTAGCCCAATGGATGCCCTGAGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
 Db 1603 CCTAGCCCAATGGATGCCCTGAGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 1662

QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTGTTAAAGTTTGTCTCTTCCAGAAAT 300
 Db 1663 GCTACTCTCTTTGGACCTGATCTTAACTCTCTTGTGTTAAAGTTTGTCTCTTCCAGAAAT 1722

QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGACGTCCATGACTAA 360
 Db 1723 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGACGTCCCAAGACTAA 1770

QY 361 AATCTACCTGGACCCCTGGACCGGCTCTGATGCTCTGATGCTCTGATGTTAATGACATGGA 420
 Db 1771 GATCTACCGCAGACCCCTGGACCGGCTCTGATGCTCTGATGCTGATGTTAATGACATCAA 1830

QY 421 AGTCACCTCTCCGAGGAATCTCACTGACAGACCCCTACTACACTTCACTCAATTCAGTAGG 480
 Db 1831 AGGCACCTCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890

QY 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTTCTCTGTTGAGA 540
 Db 1891 AAGCAGTTAGAGGTTGTTGCGCAACCTCCCAACAGCAGTACTGTTTCTCTGTTGAGA 1950

QY 541 GGGTGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGATCCCAAGCTT 600
 Db 1951 TGGGGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGATCCCAAGCTT 2010

QY 601 ANCTGGAGAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACTAGCTACACACCGG 660
 Db 2011 AGCTGGGAAGGTGACCACTCCACCTTTAAACACGGGGCTTGCACTAGCTACACACCTG 2070

QY 661 ACCAATCAGAGACTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
 Db 2071 ACCAATCAGAGACTCACTAAATGCTAATTAGGCAAAAGACAGAGGTAAAGCAATAGCC 2130

QY 721 AATCATCTATTGCTGAGAGCAGCGGGAGCAAGGATGGGATATAAATCAGGCA 780
 Db 2131 AATCATCTATTGCTGAGAGCAGCGAGGAGCAATGATCGGGATATAAATCAGGCA 2190

QY 781 TTCAAGCCAGCACAGCAACCCCTTTGGTCCCTCCCATCTGATGGGAGCTCTGTTT 840
 Db 2191 TTCAAGCCAGCACAGCAACCCCTTTGGTCCCTCCCTCTCTGATGGGAGCTCTGTTT 2250

QY 841 CACTTATTTCACTCTATTAAATCATGCAACTGCACCTCTCTGCTCGTGTGTTTATGG 900
 Db 2251 CATGCTATTTCACTCTATTAAATCTTGCACTGCACTCTCTCTGCTCAATGTTTCTACGG 2310

QY 901 CTCAGCTGAGCTTTGTTCCGCAATCAACACTGCTGTTTGGCCACCGTCAACACCGCT 960
 Db 2311 CTTGAGCTGAGCTTTGCTCGCCATCAACACTGCTGTTTGGCCCAACCGTCAACACCGCT 2370

QY 961 GCTGACTTCCATCCCTTTGATCCAGAGAGTCTCACTGCTGCTCTGATCCAGCAGGTT 1020
 Db 2371 GCTGACTTCCATCCCTTTGATCATGCAAGGTTGCTGCTGCTGATCCAGCAGGTT 2430

QY 1021 ACCATTGGCACTCCCGATCAGGCTAAGGCTTGGCACTTCTCTGCTGATGCTTAAGTCC 1080
 Db 2431 ACCATTGGCGCTCCCAATCGGCTAAGGCTTGGCACTTCTCTGCTGATGCTTAAGTCC 2490

QY 1081 TGGGTTTCTCTTAATAGAACTGAACACCTGGTCACTGGGTTCCATGGTCTCTTCCATGAC 1140
 Db 2491 TGGGTTTCTCTTAATAGAACTGAACACCTAGTCACTGGGTTCCATGGTCTCTTCTGTGAC 2550

QY 1141 CCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTTCCATTCCTTGGTA 1200
 Db 2551 CCACAGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGGTTCCATTCCTT-GAA 2609

QY 1201 TCTGTGAGCCCAAGAACCCAGGTCAGAGAANGTCAGGCTTGCACCACTTTGGGAAG 1257
 Db 2610 TCCATAAGCCCAAGAACCCAGGTCAGAGAACACAGGCTTGCACCACTTTGGGAG 2666

RESULT 15
 AF520495S2 2694 bp DNA linear PRI 11-FEB-2004
 LOCUS Homo sapiens individual 24 allele B, envelope glycoprotein gene,
 DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
 ACCESSION AF520496
 VERSION AF520496.1 GI:33410960
 KEYWORDS
 SEGMENT
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
 Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
 TITLE The endogenous retroviral locus ERVW1 is a bona fide gene involved
 in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBMED 14757826
 REFERENCE 2 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O. and Oriol,G.
 DIRECT SUBMISSION
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
 d'Italie, Lyon 69364 cedex 07, France

FEATURES
 source
 1..2694
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="individual 24 allele B"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-q22"
 /sex="female"
 /cell_type="PBMC's"
 /note="Asian
 endogenous_virus: HERV-W"
 misc_feature
 43..57
 /note="splice acceptor site"
 CDS
 284..1900
 /note="syncytin"
 /product="envelope glycoprotein"
 /protein_id="AAQ1570.1"
 /db_xref="GI:33410961"
 /translation="MALPYHIEFTVLLPSFTLTAPPCCRCMTSSSPYQEFILWRMQR
 GNIDAPSVRLSKGTPPTFAHTHPRNCVHSATLCMHANTHYTKWNPSCPGILG
 TVCWYFTQMGSDGGVQDOAREKHVKEVISQLTAVHGTSSPYKGLDLSKLHETLRT
 HTPLVSLFNTLTGLHESVQAQNPNCWICLPLNFRPYVSIIPVEQNNFSTENITTSV
 YRLGSLSSMEITCFLSFLVPMPTIYTEDQDLYSVISKPRNKRVPILDFVIGAGVIGALG
 TGGITTTSTQFYKLSQBLNGDMERVDLSVTLQDLSLAALVQLNRALDLTLAE
 RGCTCLFGECCYVNOGSIIVTEKVEIRDRIOREBELRNTGPMGLLSQWMPWTLIP
 FLGPLAAIILLLLFGPCIFNLAVNFVSSRIEAVKLQMBPKMQSKTKIYRRPLDRPAS
 RSDVNDIKGTPPEEISAAQPLRPNSAGSS"
 misc_feature
 1705..1720
 /note="splice acceptor site"
 1959..2694
 LTR

ORIGIN		Query Match 77.9%; Score 1035.2; DB 9; Length 2694;	
		Best Local Similarity 90.1%; Pred. No. 3e-304;	
		Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;	
Qy	1	TCAAATCGAAGAGCTTTAGACTGCTTAACCGCCAAAGAGAGGGGACCTGTTTATTTTT	60
Db	1423	TCAAATCGAAGAGCTTTAGACTGCTTAACCGCTGAAAGAGAGGGGAACTGTTTATTTTT	1482
Qy	61	AGGGGAAGATGCTGTAGTATGTTAAATCAATCTCGGAATCAATCTAGAGAAGTTAAAGA	120
Db	1493	AGGGGAAGATGCTGTATTTATGTTAATCAATCCGAATCGTCACTGAGAAGTTAAAGA	1542
Qy	121	AATTTGAGATCGAATATATATAGAGCAGAGAGACCTTCAAAACACTGACCCCTGGGCGCT	180
Db	1543	AATTCGAGATCGAATACAACTAGAGCAGAGAGAGCTTCGAAACACTGGACCCCTGGGCGCT	1602
Qy	191	CCTCAGCAATGGATGCCCTCGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT	240
Db	1603	CCTCAGCAATGGATGCCCTCGATTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT	1662
Qy	241	TTTACTCTCTTTTGACCCCTGATCTTCAACTCTCTTTAAAGTTTGTCTTCTCCAGAAAT	300
Db	1663	GCTACTCTCTTTTGACCCCTGATCTTAACTCTCTTTAACTTTGTTCTCTCCAGAAAT	1722
Qy	301	TGAAGCTGAAGCTACAAATAGTTCTTCAATGGAAACCCAGATGCAAGTCCATGACTAA	360
Db	1723	CGAAGCTGTAATACTA-----CAAAATGGAGCCCAAGATGCAAGTCCCAAGACTAA	1770
Qy	361	AATCTACCGTGGACCCCTGGACCGCGCTGTAGACTATGCTGTGATGTTAATGACATGGA	420
Db	1771	GATCTACCGCAGACCCCTGGACCGCGCTGTAGCCACGATCTGATGTTAATGACATCAA	1830
Qy	421	AGTCAACCCCTCCGAGGAAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGG	480
Db	1831	AGGCACCCCTCTGTAGGAAATCTCAGCTGCAACACCTCTACTAGCCCCCAATTCAGCAGG	1890
Qy	481	AAGCAGTTAGAGCAGTTGTAGCCAACTCTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540
Db	1891	AAGCAGTTAGAGCGGTCTCGGCCAACCTCTCCCAACAGCAGTACTTGGTTTTCTGTTGAGA	1950
Qy	541	GGTGGAGTCTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT	600
Db	1951	TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCT	2010
Qy	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTTGCCAATTTAGCTCACACCG	660
Db	2011	AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG	2070
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGTAAAGCATAGCC	720
Db	2071	ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAAACAGGAGTAAAGAAATAGCC	2130
Qy	721	AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780
Db	2131	AATCATCTATTGCTGAGAGCAGCAGGAGGGAATGATCGGATATAAACCAGATC	2190
Qy	781	TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATGTTATGGAGAGCTGTGTTTT	840
Db	2191	TTCGAGCGGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGAGAGCTGTGTTTT	2250
Qy	841	CACCTATTTCACCTATTAAATCATGCACTGCACTCTCTGTCGCTGTTTTTTATGG	900
Db	2251	CATGCTATTTCACCTATTAAATCTTGCAACTGCACTCTTCTGTCGCAATGTTTCTACGG	2310
Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCT	960
Db	2311	CTTGAGCTGAGCTTTTCGCTCGCCATCCACCACTGCTGTTTGGCCGCCACCGCAGACCCGCC	2370
Qy	961	GCTGATCTCCATCCCTTTGGATCCAGCAGAGTGTCCATGTGCTCTCTGATCCAGCGAGGT	1020
Db	2371	GCTGACTCCCATCCCTCTGGATCATGCGGGGTGTCGCTGCTCTCTGATCCAGCGAGGC	2430

Qy	1021	ACCATTTGCCACTCCCGATCAGGCTAAAGGCTTGCCATTGTTCTCTGCATGCTAAGTGCC	1080
Db	2431	ACGCATTGGCGCTCCCAATCGGCTAAAGGCTTGCCATTGTTCTCTGCATGCTAAGTGCC	2490
Qy	1081	TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTTCCATGAC	1140
Db	2491	TGGGTTTCATCTTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGAC	2550
Qy	1141	CCACGGCTTCTAATAGAGCTATAAACA	1200
Db	2551	CCACAGCTTCTAATAGAGCTATAAACA	2609
Qy	1201	TCTGTGAGGCAAGAACCCAGGTCAGAGAANGTGAGGCTTGCCACCAATTTGGGAAG	1257
Db	2610	TCCATAAGGCCAAGAACCCAGGTCAGAGAACACGAGGCTTGCCACCACTTTGGGAG	2666
RESULT 16			
AF520497S2			
LOCUS			
DEFINITION Homo sapiens individual 30 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.			
ACCESSION AF520498			
VERSION AF520498.1 GI:33410964			
KEYWORDS			
SEGMENT			
SOURCE			
ORGANISM			
2 of 2			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 2694)			
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.			
The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology			
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)			
14757826			
2 (bases 1 to 2694)			
Mallet,F., Bouton,O. and Oriol,G.			
Direct Submission			
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France			
Location/Qualifiers			
1..2694			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/isolate="individual 30 allele A"			
/db_xref="taxon:9606"			
/chromosome="7"			
/map="7q21-q22"			
/sex="female"			
/cell_type="PBMC's"			
/note="Caucasian			
endogenous_virus: HERV-W			
43..57			
/note="splice acceptor site"			
284..1900			
/note="syncytin"			
/codon_start=1			
/product="envelope glycoprotein"			
/protein_id="AAQ17571.1"			
/db_xref="GI:33410965"			
/translation="MALPYHFLFTVLLPSFTLTAPPBCMTSSSPVQBELMRMRQRRP			
GNIDAPSRLSLKGTPTFTATHMPNRCYHSATLCMANTHYWTGKINPSCPLGLV			
TVCWYTFQTMDSGGVQDOAREKHVKVISQLTRVHGTSSTPYKGLDLSKHLRT			
HTRLVSLFNTLTGLHEVSAQNPTNCWICLPNFRPVYSPVPEQWNNFSTEINTSV			
LVGPLVSNLSITHTSNLTCKVSNITVTNNSQCIRWTPPTQIVLPSGIFPVCGTSA			
VRLNGSSBSMCFSLFLVPMTIVTEODLYSYVTSKPNKRVPILPFLVIGAGVLGALG			
TGIGITTSQFYKLSQELNGDMERVADSLVTLQDQLNSLAAVVLNRRALDLLTAE			
RGCTCLPGEBCCYVNGQSLVTEKVKIEIDRIORRABELNRTGMLLSOMPKILP			
FLGPLAAIILLFLFGPCIFNLLVNFVSSRIEAVKLQWEPKMQSKTKIYTRPDRPAS			

RSDVNDIKGTPPEISAAQPLLRPNAGSS"		
misc_feature	1705..1720	
/note="splice acceptor site"		
LTR	1959..2694	
ORIGIN		
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;		
Best Local Similarity 90.1%; Pred. No. 3e-304;		
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;		
QY	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTTATTTT 60
DB	1423	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTT 1482
QY	61	AGGGGAAGATGCTGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA 120
DB	1483	AGGGGAAGATGCTGTATTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 1542
QY	121	AATTTGAGATCGAATATAATGTAGACGAGGAGCTTCAAAACACTGCAACCTTGGGGCCT 180
DB	1543	AATTCGAGATCGAATACAAGTAGACGAGGAGCTTCGAAACACTGGACCTTGGGGCCT 1602
QY	181	CCTCAGCCAAATGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATAT 240
DB	1603	CCTCAGCCAAATGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATAT 1662
QY	241	TTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGTAGTTTGTCTCTTCCAGAA 300
DB	1663	GCTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGTAGTTTGTCTCTTCCAGAA 1722
QY	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATCGATCCATGACTAA 360
DB	1723	CGAAGCTGTAAAGCTA-----CAATGGAAGCCCAAGATCGATCCAGACTAA 1770
QY	361	AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 420
DB	1771	GATCTACCGGAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830
QY	421	AGTCACCCCTCCCGAGGAATCTCAACTGACGACAAACCCCTACTACATCCAAATCAGTAG 480
DB	1831	AGGCACCCCTCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCACGAGG 1890
QY	481	AAGCAGTTAGACAGTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
DB	1891	AAGCAGTTAGACAGTTGTCGCGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 1950
QY	541	GGGTGAGCTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGAATCCCNAAAGCCT 600
DB	1951	TGGGGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGCTGACTAAGAATCCCTAAGCCT 2010
QY	601	ANCTGGGAAGGTGACCGCATCTTCTTAAACATAGGGGCTTGCACTAGCTCACACCCG 660
DB	2011	AGCTGGGAAGGTGACCATCTCCACCTTTAAACACGGGGCTTGCACTAGCTCACACCTG 2070
QY	661	ACCAATCAGAGCTCCTAAATATGCTAATCAGGCAAAACAGAGGTAAGCAATAGCC 720
DB	2071	ACCAATCAGAGCTCCTAAATATGCTAATAGGCAAAAGCAGAGGTAAGCAATAGCC 2130
QY	721	AATCATCTATTGCTGAGAGCACAGCGGGAGGACAGGATGCGGATATAAACTCAGGCA 780
DB	2131	AATCATCTATTGCTGAGAGCACAGCGAGGAGGACAAATGATCGGATATAAAACCCAGTC 2190
QY	781	TTCAAGCGCAACAGCAACCCCTTTGGTCCCTCCCTCCCTTTGATGGAGCTCTGTTT 840
DB	2191	TTCAAGCGCGCAACAGCAACCCCTTTGGTCCCTCCCTCCCTTTGATGGAGCTCTGTTT 2250
QY	841	CACCTCTATTCTCTATTAAATCATGCACTCTCTCTGCTCGCTGCTTTTATGG 900
DB	2251	CATGCTATTCTCTATTAAATCTTGCACACTGCATCTTCTGCTCAATGTTTCTTACGG 2310
QY	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGCTCACAGACCGCT 960
DB	2311	CTTGAGCTGAGCTTTGCTCGCCATCCACACTGCTGTTTGGCCACCGCCACAGACCGCC 2370

QY	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAGGT 1020
DB	2371	GCTGACTTCCATCCCTCTGGATCATGCGAGGTGTCCGCTGTCTCTGATCCAGCGAGGC 2430
QY	1021	ACCCATTGCCACTCCCGATCAGGCTAAAGGCTTGGCAATTTCTTGCATGGCTTAAGTGCC 1080
DB	2431	ACCCATTGCCGCTCCCAATCGGCTTAAAGGCTTGGCAATTTCTTGCATGGCTTAAGTGCC 2490
QY	1081	TGGGTTTCTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGTTCTCTTCCATGAC 1140
DB	2491	TGGGTTTCTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGTTCTCTTCTGAC 2550
QY	1141	CCACCGCTTCTTAATAGAGCTATAACACTCACCGCATGCGCCCAAGATTTCCATTTCTTGGTA 1200
DB	2551	CCACAGCTTCTTAATAGAGCTATAACACTCACCGCATGCGCCCAAGATTTCCATTTCTT-GAA 2609
QY	1201	TCTGTGAGGCCAAGAACCCAGGTCAGAGAANGTGAAGCTTGCACCATTTGGGAAG 1257
DB	2610	TCCATAAGGCCAAGAACCCAGGTCAGAGAACACGAGGCTTGCACCATTTGGGAG 2666
RESULT 17		
AF520503S2		
LOCUS		
DEFINITION		
Homo sapiens individual 34 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.		
ACCESSION		
AF520504		
VERSION		
AF520504.1 GI:33410976		
KEYWORDS		
2 of 2		
SOURCE		
Homo sapiens (human)		
ORGANISM		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE		
1 (bases 1 to 2694)		
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.		
The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology		
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
JOURNAL		
PUBMED		
14757826		
REFERENCE		
2 (bases 1 to 2694)		
Mallet,F., Bouton,O. and Oriol,G.		
Direct Submission		
TITILE		
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France		
JOURNAL		
FEATURES		
Location/Qualifiers		
1..2694		
/organism="Homo sapiens"		
/mol_type="genomic DNA"		
/isolate="individual 34 allele B"		
/db_xref="taxon:9606"		
/chromosomes="7"		
/map="7q21-q22"		
/sex="male"		
/cell_type="PBMC's"		
/note="endogenous_virus: HERV-W"		
43..57		
/note="splice acceptor site"		
284..1900		
/note="syncytin"		
/codon_start=1		
/product="envelope glycoprotein"		
/protein_id="AAQ17574.1"		
/db_xref="GI:3341097"		
/translation="MALPYHIEFTVLLPSFTLTAPPCCRCMTSSPYOEFLWRMORP GNIDAPSVRLSKGPTTFAHTHMPRNCVHSATLCWHANTHYWTGMINPSCPGGLGV TYCWTFTQTGMSDGGVQDQAREKHKEVISQLTRVHCTSSPYKGLDLSKHETLRT HYRLVFNNTLTGLHVSQAQNPNCICLPLNFRPVSIPVPEQWNNESTEINTTSV LVGLVSNLEITHLTCVKFESNTTYSQCIRWVTPTQIVCLPSGIFPVCVGTSA YRLNGSSSMCMFLSFLVPMPTIYTEQDLYSYVISKPRNRPVILFPVIGVILGALG		
misc_feature		
CDS		


```

TGIGGITTSTQFYKLSQELNGDMERVADSLVTLQDQNSLAAVVLQNRALDLTAE
RGGCLFLGECGGYVNSGIVTEKVEIRDRIOORABEELNRTGFWLLSOMFWIIP
PLGFLAAIILLLEFGPCIFNLNVFVSRIEAVLQMEPKMQSKTKIYRRLDRPAS
RSDVNDIKGTTPPEISAAQPLLRPNISAGSS"
1705..1720
/misc_feature
/notes="splice acceptor site"
1959..2694
ORIGIN
Query Match      77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGAGGGGAACCTGTTATTTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCTGAAAGAGAGGGGAACCTGTTATTTTT 1482
QY 61 AGGGAAGAACTCTTAGTAGTATTAATCAATCTGGAATCAATCTAGAGAAGTTAAAGA 120
Db 1483 AGGGAAGAACTCTTAGTAGTATTAATCAATCTGGAATCAATCTAGAGAAGTTAAAGA 1542
QY 121 AATTGTAGATCGAATATAATGTAGAGCAGAGGACCTTCAAAACACTGCACCCCTGGGCCT 180
Db 1543 AATTGTAGATCGAATATAACAGGTAGAGCAGAGGAGCTTCGAAACACTGCACCCCTGGGCCT 1602
QY 181 CCTAGCAATGATGACCCCTGGACTCTCCCTCTTAGGACTCTAGCAGCTATAATATT 240
Db 1603 CCTAGCAATGATGACCCCTGGACTCTCCCTCTTAGGACTCTAGCAGCTATAATATT 1662
QY 241 TTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTGTTAAAGTTGTTCTCTTCAGAAAT 300
Db 1663 GCTACTCTCTTTGGACCCCTGATCTTTAACTCTCTGTTAAAGTTGTTCTCTTCAGAAAT 1722
QY 301 TGAAGCTGTAAAGCTACAATAGTTCTTCAATGGAACCCAGATGCGAGTCCATGACTAA 360
Db 1723 CGAAGCTGTAAAGCTA-----CAATGAGGCCAAGATGCGAGTCCAAGACTAA 1770
QY 361 AATCTACGTGAGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 1771 GATCTACGCGACAGCCCTGGACCGGCTGTAGCCAGCATGTGATGTTAATGACATCAA 1830
QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAATTCAGTAGG 480
Db 1831 AGGCACCCCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGACAGTTGTGACCCAACTCTCCCAAAGTACTCTGCGGTTTCTCTGTTAGA 540
Db 1891 AAGCAGTTAGACGGGTGCTGCGCCAACTCTCCCAAAGTACTCTGCGGTTTCTCTGTTAGA 1950
QY 541 GGGTGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT 600
Db 1951 TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 2011 AGCTGGGAAGGTGACCAATCCACTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2070
QY 661 ACCAATCAGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
Db 2131 AATCATCTATTGCTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCCAGTCT 2190
QY 781 TTCAAGCCAGCAACACCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
Db 2191 TTCAAGCCAGCAACGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 2250
QY 841 CACTCTATTCTCTATTAAATCATGCAACTGCACTCTCTGGTCCGTTTATATGG 900
Db 2251 CATGCTATTCTCTATTAAATCTTTGCAACTGCACTCTCTGGTCCCTCTTCTTACCG 2310
```

```

QY 901 CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTGCCACCGCTCAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTGCCACCGCTCAGACCCGCT 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGGT 1020
Db 2371 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGGT 2430
QY 1021 ACCATTGACCACTCCCGATCAGGCTAAAGGCTTGCCATTTGCTCTGATGCTAAAGTGCC 1080
Db 2431 ACCATTGACCACTCCCGATCAGGCTAAAGGCTTGCCATTTGCTCTGATGCTAAAGTGCC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTTCATCTGGGTTCCATGTTCTTCCATGAC 1140
Db 2491 TGGGTTTGTCTTAATAGAACTGAACACTGGTTCATCTGGGTTCCATGTTCTTCTGTCAC 2550
QY 1141 CCAGCGGTTCTTAATAGAGCTATAACTCACTACCGCATGGCCAGATTCATCTCTTGGTA 1200
Db 2551 CCAGCGGTTCTTAATAGAGCTATAACTCACTACCGCATGGCCAGATTCATCTCTTGGTA 2609
QY 1201 TCTGTGAGGCCAAGAACCCAGGTCAGAGGCTTGAGGCTTGCCACCATTTGGGAAG 1257
Db 2610 TCCATAAGGCCAAGAACCCAGGTCAGAGGCTTGAGGCTTGCCACCATTTGGGAAG 2666

RESULT 18
AF52050982 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 45 allele A, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520510
VERSION AF520510.1 GI:33410988
KEYWORDS
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
Location/Qualifiers
source
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 45 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMc's"
/notes="endogenous_virus: HERV-W"
43..57
/misc_feature /note="splice acceptor site"
284..1900
CDS
/note="syncytin"
/product="envelope glycoprotein"
/protein_id="AAQ17577.1"
/db_xref="GI:33410989"
/tranlation="MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPQYOEFLWRMQR
PNIDAPSYSRLSKGTPPTTAHTMPNRCVHSATLCMHANTHPCVSPKQINPSCPGGLGV
TVCWTYFTIQMSDGGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT
```



```

HTRLVSLFNTLTGLHEVSAQNPTNCWICLPLNFRPYVSIIPVPEOMNFSINTTSS
LVGPLVSNLEIHTHSNLCVAFKSNNTYITNSOCIRWTPPTQIVCLPSGIFFVCGTSA
YRCLNGSSESMCFSLFVPPMTIYTEQDLYSVIVISKPENKVPVLPFVIGAGVLGALG
TGIGDITSTQFYKLSQELQDMERVADSLVTLQDQNSLAADVVLQNRALDILLTAE
RGCTFLGEBECYVNGSGIVTEKVRIDRIORRABELRNTPGWLLSOWMPWILP
FLGPLAAIILLLFPFCIFNLVNFVSSRIEAVKLQMEFQMSKTKIYRRPLDRPAS
RSDVNDIKGTPEEBEISAQPLLRPNASGS"
misc_feature
1705..1720
/note="splice acceptor site"
LTR
1959..2694
ORIGIN
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGCTAAACCCGCAAGAGAGGGGAAACCTGTTTATTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTTGCTAAACCCGCTGAAGAGGGGAAACCTGTTTATTTT 1482
QY 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCATCTAGAGAAAGTTAAAGA 120
Db 1483 AGGGGAAGAAATGCTGTATATGTTAATCAATCCGGAATCGTCACTGAGAAAAGTTAAAGA 1542
QY 121 AATTTGAGATCGAATATAATGATAGACGAGAGACCTTCAAAACACTGACCCCTGGGGCT 180
Db 1543 AATTCGAGATCGAATACAACGTAGACGAGAGAGCTTCGAAACACTGGACCCCTGGGGCT 1602
QY 181 CCTCAGCCAAATGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 1603 CCTCAGCCAAATGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1662
QY 241 TTTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGTAGTTTGTCTCTTCCAGAA 300
Db 1663 GCTACTCTCTTTGGACCTGTATCTTAACTCTCTTGTAACTCTTGTCTCTTCCAGAA 1722
QY 301 TGAAGCTGTAAGACTACAATAATGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA 360
Db 1723 CGAAGCTGTAAGACTA-----CAATGGAGCCCAAGATGCAGTCCAGACTAA 1770
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 420
Db 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830
QY 421 AGTCACCCCTCCCGAGAAATCTCAACTGACACACCCCTACTACACTCCAAATTCAGTAG 480
Db 1831 AGGCACCCCTCTCTGAGGAAATCTCAGTGCACAACTCTACTACGCCCCCAATTCAGCAG 1890
QY 481 AAGCAGTTAGACAGTGTGTCAGCAACCTCCCAACAGTACTTTGGGTTTCTCTGTTGAGA 540
Db 1891 AAGCAGTTAGACGGTGTGTCAGCAACCTCCCAACAGCAGTACTTTCTCTGTTGAGA 1950
QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 1951 TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGGCACTAGCTAGCTACACCC 660
Db 2011 AGCTGGGAAGGTGACCACTCCACCTTTAAACACGGGGCTTGGCACTAGCTAGCTACACCT 2070
QY 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGAGTAAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGCTCACTAAATGCTTAATAGGCAAAAGCAGAGAGTAAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGAGCAGCGGGGAAGGACCAAGGATTTGGGATATAAACTCAGGCA 780
Db 2131 AATCATCTATTGCTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAAACCAAGTC 2190
QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
Db 2191 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTT 2250
QY 841 CACTCTATTTCACCTCTTAATAATCATGCACTCTTCTCGTCCGTGTTTTTATGG 900
```

```

Db 2251 CATGCTATTTCACCTCTATTAAATCTTGCAACTGCACCTCTCTGGTCCATGTTTCTTAGCG 2310
QY 901 CTCAGCTGAGCTTTGTTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTGCTGCTGCCATCCACCACTGCTGTTTGGCCGACCGGACAGACCGGC 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGGAGGT 1020
Db 2371 GCTGACTCCCATCCCTCTGGATCATGCAGGGTGTCCGCTGTGCTCTGATCCAGGAGGC 2430
QY 1021 ACCCATTCGCCATCCCGATCAGGCTAAAGGCTTGCCATGTTCTTCGATGGTAAAGTGC 1080
Db 2431 ACCCATTCGCCCTCCCAATCGGGCTAAAGGCTTGCCATGTTTCTTCGATGGTAAAGTGC 2490
QY 1081 TGGGTTTCTCTTAATAGAACTGAAACACTGCTCAGTGGTTCCTCATGTTCTCTTCCATGAC 1140
Db 2491 TGGGTTTCTCTTAATGAGCTGAAACACTAGTACATGAGGTTTCCATGTTCTCTTCTGTGAC 2550
QY 1141 CCACGGCTTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGGTA 1200
Db 2551 CCACAGCTTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTT-GAA 2609
QY 1201 TCTGTGAGGCCCAAGAAACCCAGGTCAGAGAANGTGAGGCTTCCACCATTTGGGAAG 1257
Db 2610 TCCATAAGGCCCAAGAAACCCAGGTCAGAGAACACGAGGCTTCCACCATTTGGGAG 2666

RESULT 19
AF520513S2 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 49 allele A, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520514
VERSION AF520514.1 GI:33410996
KEYWORDS 2 of 2
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
Location/Qualifiers
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 49 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="Asian
endogenous_virus: HERV-W"
43..57
/note="splice acceptor site"
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17579.1"
misc_feature
284..1900
CDS
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17579.1"
```

```
/db_xref="GI:33410997"
/translation="MALPYHIFLFTVLLPSTLUTAPPCKMTSSPYQFLWRMORP
TNDAPYSRLSKGTPFTFAHTMPNCHYSAITLCHNHTYWTGMNPSCPGILGV
GVNWTYFTQTMGSGGGVQOAREKHKEVISOLTRVHGTSSTPYKGLDLSKLHETLRT
HTRLVSLPNTLTGLHVSQAONPNWCILPLNFRPVSVIPVPEOWNFSTENITTSV
LVGPLVSNLEITHFSLNLTVCVFNSTYTTTNSQICRWVTPPTQIVCLPSGIFVCGTSA
YRCLVNSESMECHFSFLVFPWMTIYEODLYSVISKPRNKRVPILFPVIGAGVIGALG
TGIGITTSQFYFKJQLQDNGDMERVADSLVTLQDOLNSLAADVLRNRRALDLILLT
RGCTCLFLGBECCYVNNQSGIVTEKVEIKRDIRQRAABELRNTGPMGLLSQWMPWILP
PLGPLAAIILLLPLGPCIFNLLVNFVSSRIEAVKLQMEPKNQSKTKIYRPLDRPASP
RSDVNDIKGTPPEISAAQPLLRNSAGSS"
misc_feature
1705..1720
/note="splice acceptor site"
LTR
1959..2694
ORIGIN
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 113; Conservative 0; Mismatches 11; Indels 13; Gaps 2;

QY 1 TC AAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGAACTGTTTATTTT 60
Db 1423 TC AAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTT 1482
QY 61 AGGGAAGAATGCTGTTAGTAGTATGTTAATCAATCTGGAATCAATCTGAGAAAGTTAAAGA 120
Db 1483 AGGGAAGAATGCTGTTATGTTAATCAATCTGGAATCAATCTGAGAAAGTTAAAGA 1542
QY 121 AATTGAGATCGAATATATGATAGACGAGAGACCTTCAAACACACTGCACCCCTGGGGCCT 180
Db 1543 AATTCGAGATCGAATACACGCTAGACGAGAGGAGCTTCGAACACACTGCAGCCCTGGGGCCT 1602
QY 181 CCTCAGCCAATGATGATGCTGCTGCTCCCTTTCTAGGACTCTAGCAGCTATAATTT 240
Db 1603 CCTCAGCCAATGATGATGCTGCTGCTCCCTTTCTAGGACTCTAGCAGCTATAATTT 1662
QY 241 TTTACTCTCTTTGAGACCTGATCTTCAACTCTCTGTTAAGTTTGTCTTCTCCAGAAAT 300
Db 1663 GCTACTCTCTTTGAGACCTGATCTTAAACCTCTCTGTTAAGTTTGTCTTCTCCAGAAAT 1722
QY 301 TCAAGCTGTAAAGCTACAATAAGTTCTTCAAAATGGAACCCCAAGATGCAAGTCCATGACTAA 360
Db 1723 CGAAGCTGTAAACTA-----CAAAATGAGCCCAAGATGCAAGTCCAAGACTAA 1770
QY 361 AATCTACCGTGAACCCCTGGACCGGCTGCTAGACTGATGCTCTGATGATTAATGACATGA 420
Db 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGACCGATCTGATGATTAATGACATCAA 1830
QY 421 AGTCACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAATTCAGTAGG 480
Db 1831 AGGCACCCCTCTGAGGAATCTCAGCTGCACAACTTACTACGCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGACAGTTGTGACCAACCTCCCAACAGTACTTTGGGTTTCTGTTGAGA 540
Db 1891 AAGCAGTTAGACAGTTGTGCGGCAACCTCCCAACAGCAGTACTTTCTGTTGAGA 1950
QY 541 GGGTGACTGAGACAGACGACTAGCTGGATTTCTAGGCTGACTAAGATCCCNAGCCT 600
Db 1951 TGGGGGACTGAGACAGACGACTAGCTGGATTTCTAGGCTGACTAAGATTCCTTAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCGG 660
Db 2011 AGCTGGGAAGGTGACCAATCCACCTTTAAACAGGGGCTTGCAACTTAGCTCACACCTG 2070
QY 661 ACCAATCAGACAGCTCACTAAATATGCTAATTCAGGCAAAAACAGGAGGTTAAGCAATAGCC 720
Db 2071 ACCAATCAGACAGCTCACTAAATATGCTAATTCAGGCAAAAACAGGAGGTTAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTCAGAGCAGCGGGAAGGACAGGATTTGGGATATAAACTCAGGCA 780
Db 2131 AATCATCTATTGCTCAGAGCAGCAGGAGGAGCAATGATCGGATATATAACCCNAGTC 2190
QY 781 TTCAAGCCAGCAACAGCAACACCCCTTTGGGTCCCCCTCCCAATTTGATGGAGCTCTGTTTT 840
```

```
Db 2191 TTCGACCGGCAACGCAACCCCTTTGGTCCCCTCCCTTTGTATGGAGCTCTGTTTT 2250
QY 841 CACTCTATTCTACTCTATTAAATCATGCAACTGCACCTCTCTGGTCCGCTTTTTTATGG 900
Db 2251 CATGCTATTCTACTCTATTAAATCTTTGCAACTGCACCTCTCTGGTCCATGTTTCTTACGG 2310
QY 901 CTCAACTCAGCTTTTGTTCGCGCATCCACACTGCTGTTTGGCACCCTGTCACAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTCGCTCGCCATCCACACTGCTGTTTGGCGCCACCGCAGACCCGCC 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGTCTCTGATCCAGCGAGGT 1020
Db 2371 GCTGACTCCCATCCCTCTGATCATGAGGGTGTCCGCTGTCTCTGATCCAGCGAGGC 2430
QY 1021 ACCATTTGCCACTCCCGATCAGCTAAAGCTTGCATTTGCTCCTCATGCTGAAGTGCC 1080
Db 2431 ACCATTTGGCGCTCCCAATGGGCTAAAGCTTTGCCATTTGTTCTCGATGCTAAGTGCC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACACTGTGTCTACTGGGTTCCATGTTCTTTCATGAC 1140
Db 2491 TGGGTTTCACTTAATTTGAGCTGAACACTAGTCACTGGGTTCCATGTTCTTCTGTGAC 2550
QY 1141 CCACGGCTTCTAATAGAGCTATAAACHACTACCCGATGGCCCAAGATTCATTTCTTTGGTA 1200
Db 2551 CCACAGCTTCTAATAGAGCTATAAACHACTACCCGATGGCCCAAGGTTCCATTTCTT-GAA 2609
QY 1201 TCTGTGAGGCCAAGAACCCCGAGCTCAGAGAANGTGAGGCTTGCCACCATTTGGGAAG 1257
Db 2610 TCCATTAAGGCCAAGAACCCCGAGCTCAGAGAANGTGAGGCTTGCCACCATTTGGGAG 2666

RESULT 20
AF520515S2 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 49 allele B, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520516
VERSION AF520516.1 GI:33411000
KEYWORDS
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MOLTYPE: nucleotide
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWEL is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
DIRECT SUBMISSION
TITLE Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
JOURNAL CNRS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 49 allele B"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/feature="Asian
endogenous_virus: HERV-W"
misc_feature 43..57
/feature="splice acceptor site"
CDS 284..1900
```

```

/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17580.1"
/db_xref="GI:33411001"
/translation="MALPYHFLFTVLPSFTLAPPCCRCMTSSSPYOEFWRMRP
GNIDAPSRSLSKGTPTTAHTHPRNCYHSATLCMHANTHWYTGKMLNPSCPGLGV
TWCTYPTQDSGDDGVDQAREKHVEVTSQLTRVHGTSPPYKGLDLXKHLRT
HTRVLSNLEITHTSNLTCVKSNTYTTNSQIRWVTPPTQIVCLPSGIFVCGTSA
LVGPLVSLNLEITHTSNLTCVKSNTYTTNSQIRWVTPPTQIVCLPSGIFVCGTSA
YRLNGSSESCFTYFVPMPTIYEDQLYSVTSKPRNRPVLPFPVIGAGVLGALG
TGIGTSTGYFYKLSOBLNGMDVADSLVLODLNSLAVALQNRALDLITAE
RGCTCLFGEBCCVYNGSGIVTEKVEIRDIORAEELNRTGPGLLSOWMPWLP
FLGPLAAILLLFQPCFLNLFVSSRIEAVKLQMEPKWQSKTKIYERPLDRPAS
RSDVNDIKGTPEEISAAQPLLRLPSAGSS"
misc_feature
1705..1720
/feature_id="1705"
/note="splice acceptor site"
1959..2694
LTR
ORIGIN
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAGAGGGGAACTGTTATTTT 60
DB 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCTGAAAGAGGGGAACTGTTATTTT 1482
QY 61 AGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCATTCTAGAGAAAGTTAAAGA 120
DB 1483 AGGGAAGAAATGCTGTATATGTTAATCAATCGGAATCGTCACTAGAGAAAGTTAAAGA 1542
QY 121 AATTTGAGATCGAATATAATAGTAGAGAGAGAGACTTCAAAACACTGCACCCCTGGGGCT 180
DB 1543 AATTCGAGATCGAATACACAGTAGAGAGAGAGAGCTTCAAAACACTGCAGCCCTGGGGCT 1602
QY 181 CCTCAGCAATGAGTCCCTGGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
DB 1603 CCTCAGCAATGAGTCCCTGGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 1662
QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTCTTTGTAAGTTTGTCTCTTCCAGAA 300
DB 1663 GCTACTCTCTTTGGACCTGTATCTTAACTCTCTTTGTAAGTTTGTCTCTTCCAGAA 1722
QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCGAGATCGATCCATGCTAA 360
DB 1723 CGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCGAGATCGATCCATGCTAA 1770
QY 361 AATCTACCTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 1771 GATCTACCGAGACCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTAATGACATTGA 1830
QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAG 480
DB 1831 AGGCACCCCTCCCGAGGAATCTCAGCTGCACAAACCCCTACTACAGCCCAATTCAGTAG 1890
QY 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTTGGGTTTCTGTTGAGA 540
DB 1891 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTTGGGTTTCTGTTGAGA 1950
QY 541 GGTGAGTGTAGAGACAGACTAGTGGATTTCTTAGGCTGACTAGAAATCCCNAGCCT 600
DB 1951 TGGGGGACTGTAGAGACAGACTAGTGGATTTCTTAGGCTGACTAGAAATCCCNAGCCT 2010
QY 601 ANCTGGGAAGTGCACCGCATCTTAAACATGGGCTTGAACCTTACTGCTCACACCCG 660
DB 2011 AGCTGGGAAGTGCACACATCTCCACCTTAAACAGGGGCTTGAACCTTACTGCTCACACCTG 2070
QY 661 ACCAATCAGAGAGCTCACTAAATAGTCAATGAGCAAAACAGAGAGTAAAGCAATAGCC 720
DB 2071 ACCAATCAGAGAGCTCACTAAATAGTCAATGAGCAAAACAGAGAGTAAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGACACAGCGGAGGAGCAAGGATTGGGATATAAATCTCAGGCA 780

```

```

2131 AATCATCTATTGCTGAGAGCACAGCAGGAGGGAACAATGATCGGATATAAACCAGTC 2190
QY 781 TTCAAGCAGCAACAGCAACCCCTTTGGTCCCTCCCTCCCTATGTTATGAGGAGCTCTGTTT 840
DB 2191 TTCGAGCGGCAACGGCAACCCCTTTGGTCCCTCCCTCCCTATGTTATGAGGAGCTCTGTTT 2250
QY 841 CACTCTATTTTCACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCCTGTTTATGG 900
DB 2251 CATGCTATTTTCACTCTATTAAATCTTGAACCTGCACTCTTCTGGTCCCTGTTTATGG 2310
QY 901 CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAAGACCCGCT 960
DB 2311 CTTGAGCTGAGCTTTGCTGCTGCCATCCACCACTGCTGTTTGGCCACCGTCAAGACCCGCT 2370
QY 961 GCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGCACCTGCTCTGATCCAGCAGGT 1020
DB 2371 GCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGCACCTGCTCTGATCCAGCAGGT 2430
QY 1021 ACCCAATTCGCACTCCCGATCAGGCTAAAGGCTTGGCAATTTCTTCGATGGCTAAAGTCC 1080
DB 2431 ACCCAATTCGCGCTCCCAATCGGCTAAAGGCTTGGCAATTTCTTCGATGGCTAAAGTCC 2490
QY 1081 TGGGTTTCTCTTAATAGAGCTATAACACTCAGGATGCGGCTTCCATTTCTTCTGATGGA 1140
DB 2491 TGGGTTTCTCTTAATAGAGCTATAACACTCAGGATGCGGCTTCCATTTCTTCTGATGGA 2550
QY 1141 CCACGGCTTCTTAATAGAGCTATAACACTCAGGATGCGGCTTCCATTTCTTCTGATGGA 1200
DB 2551 CCACGGCTTCTTAATAGAGCTATAACACTCAGGATGCGGCTTCCATTTCTTCTGATGGA 2609
QY 1201 TCTGTGAGGCGCAAGAACCCAGGCTCAGAGAANGTGGGCTTCCACCACTTTGGGAG 1257
DB 2610 TCAATAGGCGCAAGAACCCAGGCTCAGAGAANGTGGGCTTCCACCACTTTGGGAG 2666

RESULT 21
AF520517S2 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 55 allele A, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520518
VERSION AF520518.1 GI:33411004
KEYWORDS
SEGMENT 2 of 2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Chevnet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 55 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"

```

[illegible]

		/sex="female"
		/cell_type="PBMC's"
		/note="endogenous_virus: HERV-W"
misc_feature	43..57	
		/note="splice acceptor site"
CDS	284..1900	
		/note="syncytin"
		/codon start=1
		/product="envelope glycoprotein"
		/protein_id="AAQ17582.1"
		/db_xref="GI:33411009"
		/translation="MALPPIFLFTVLPSFLTAPPPRCMTSSSPYOEFLWRMORP GNIDAPSYSLSGKGTFTTAHTHPNRCNTHSATLCHMANTHWTKMINPSPGGLGV TCWVYFTQGTMLDGGVQDQAREKHKEVIQSLTRVHGTSPPYKQMLDLSUKHETLRT HTRLFTNTTTLGLHEVAQNPQRCNICLPLNFRPYVSIPIVPEQNNFSTINTSV LGPLVSNLEIHTHSNLTCKVKNSTTYTNSOCIRVWTPPTQIVCLPSGIFVCGTSA YRCLGSSSESMCLFSLVPPMTLYTEQDLNVYISKPRNRPVILPFIAGVGLGALG TGLGISTTSQFYKLSBELNGDMERVASLYTLQDLNSLAAVLQNERADLLTAE RGSTCLFLGEECCYYNQSGIVTEKKEIRDEIQRABEELNTGPKLLISQWPHILP FLGLAAIITLLFLFGPCI FNLLYFVSSRIEAVKLQMEKPMQSKTKIYRRPLDRPAS RSDVNDIKGTPPEISAAOPLLRPSAGSS"
misc_feature	1705..1720	
		/note="splice acceptor site"
LTR	1959..2694	
ORIGIN		
	Query Match	77.9%; Score 1035.2; DB 9; Length 2694;
	Best Local Similarity	90.1%; Pred. No. 3e-304;
	Matches 113;	Conservative 0; Mismatches 111; Indels 13; Gaps 2;
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTTATTTT 60
Db	1423	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTT 1482
Qy	61	AGGGAGAAGTCTGTTAGTATGTTATCAATCTGGAACTATTACTAGAGAACTTAAAGA 120
Db	1483	AGGGAGAAGTCTGTTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 1542
Qy	121	AATTTGAGATCGAATATAATGTAGACAGAGGAGCTTTCAAAACACTCACCCCTGGGGCCT 180
Db	1543	AATTCGAGATCGAATACAACTGAGAGCAGAGGAGCTTCGAAACACTGGACCTGGGGCCT 1602
Qy	181	CCTCAGCAATGGATGCTCCGCTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT 240
Db	1603	CCTCAGCAATGGATGCTCCGATTCCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 1662
Qy	241	TTTACTCTCTTTGGACCGTGATCTTCAACTCTCTTGAAGTTTGCTCTCTCCAGAAT 300
Db	1663	GCTACTCTCTTTGGACCGCTGATCTTTAACCTCTCTGTTAACTTTGCTCTCTCCAGAAT 1722
Qy	301	TGAAGCTGTAAGAGCTACAAATAGTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db	1723	CGAAGCTGTAAACTTA-----CAAATGGAGCCCAAGATGAGTCCAGACTAA 1770
Qy	361	AATCTACCGTGGACCCCTGGACCGGCTGTCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db	1771	GATCTACCGCAGACCCCTGGACCGGCTGTCTAGCCACGACTCTGATGTTAATGACATCAA 1830
Qy	421	AGTCACCCCTCCGAGGAAATCTCAACTGACACACCCCTACTACACTCCAAATTCAGTAGG 480
Db	1831	AGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
Qy	481	AAGCAGTTTAGAGCAGTGTCTCAGCCAACTCCCCCAACAGTACTCTGGGTTTTCCTGTCAGA 540
Db	1891	AAGCAGTTTAGAGCGTGTCTGGCCAACTCTCCCAACAGCACTTAGGTTTTCCTGTTGAGA 1950
Qy	541	GGGTGGACTGAGACAGGACTAGCTGATTTCTTAGCTGACTAAGAAATCCCAAGCCCT 600
Db	1951	TGGGGGACTGAGACAGGACTAGCTGATTTCTTAGCTGACTAAGAAATCCCTAAGCCT 2010
Qy	601	ANCTGGGAAGTGACCGCATCTTTAAACATGGGGCTTGGAACTTAGCTTACACCCG 660
Db	2011	AGCTGGGAAGTGACCAATCCACTTTTAAACACGGGGCTTGGAACTTAGCTTACACCTG 2070

[illegible]

1831	Db	1771	GATCTACCGCAGACCCCTCGAGCGGCTGCTAGCCCAACGATCTGATGTGTTTAATGACATCA	1831
421	Qy	421	AGTCACCCCTCCGGAGGAAATCTCAACTGCAACACCCCTACTACTACATCCCAATTCAGTAGG	480
1831	Db	1831	AGGCACCCCTCTCTGAGGAAATCTCAGCTGCACAACTCTACTAGCGCCCAATTCAGCAGG	1890
481	Qy	481	AAGCAGTTAGAGCAGTTGTCAAGCAACCTCCCCAAACAGTACTTTGGGTTTTTCCTGTTGAGA	540
1891	Db	1891	AAGCAGTTAGAGCGTCTGTGGCCAACTCCCAACAGACATTTAGTTTTCTCTGTTGAGA	1950
541	Qy	541	GGGTGAGCTTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAGCCT	600
1951	Db	1951	TGGGGGACTTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTTAAGCCT	2010
601	Qy	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCNACTTAGCTCACACCG	660
2011	Db	2011	AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGAACCTTAGCTCACACCTG	2070
661	Qy	661	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGCGCAAAACAGAGAGTAAAGCAATAGCC	720
2071	Db	2071	ACCAATCAGAGAGCTCACTAAATGCTTAATAGGCAAAACAGAGAGTAAAGCAATAGCC	2130
721	Qy	721	AATCATCTATTGCTCTGAGAGCACAGCGGGAAGGACAAGGATGGGATATAAACTCAGGCA	780
2131	Db	2131	AATCATCTATTGCTCTGAGAGCACAGCAGGAGGAGCAATGATCGGATATAAACCCAAATC	2190
781	Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTGCCCTCCCAATTGATGGGAGCTCTGTTTT	840
2191	Db	2191	TTGAGCGGCAACGGCAACCCCTTTGGGTGCCCTCCCTTTGTTATGGGAGCTCTGTTTT	2250
841	Qy	841	CACCTCTATTTCACCTCTATTAAATCATGCACTGCACCTCTCTGGTCGGTGTGTTTTATGG	900
2251	Db	2251	CATGCTCTATTTCACCTCTATTAAATCTTGCAACTGCACTCTCTGGTCCATGTTCTTACGG	2310
901	Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGCCACCGCTCAAGACCCGCT	960
2311	Db	2311	CTTGAGCTGAGCTTTGCTGCGCATCCACCACTGCTGTTTGGCCACCGCAGACCCACC	2370
961	Qy	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACGTGCTCTGATCCAGGAGGT	1020
2371	Db	2371	GCTGACTCCCATCCCTCTGATCATGAGGGTGTCCGCTGTGCTCTGATCCAGCGAGGC	2430
1021	Qy	1021	ACCATGGCACATCCCGATCAGGCTAAGGCTTGCATTTGTCCTGCATGGCTAAGTGCC	1080
2431	Db	2431	ACCCATTGCGGCTCCCAATCGGCTAAGGCTTTGCCAATGTTCTCTGCATGGCTAAGTGCC	2490
1081	Qy	1081	TGGGTTTGCTCTAATAGAACTGAAACCTGGTCACTGGGTTCATGGTTCTCTCCATGAC	1140
2491	Db	2491	TGGGTTTCATCTAATTGAGCTGAAACCTAGTCACTGGGTTCCATGGTTCTCTCTGTGAC	2550
1141	Qy	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATCCATTCCTTGGTA	1200
2551	Db	2551	CCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGGTTCCATTCCTT-GAA	2600
1201	Qy	1201	TCTGTGAGGCCAAGAACCCAGGTCAGAGAANGTAGGCTTGCACCAATTTGGGAG	1257
2610	Db	2610	TCCATAAGGCCAAGAACCCAGGTCAGAGAACACAGGCTTGCACCACTCTTGGAG	2666
RESULT 27				
AF520537S2				
LOCUS				
DEFINITION				
Homo sapiens individual 81 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.				
AF520538				
VERSION				
AF520538.1 GI:33411044				
KEYWORDS				
2 of 2				
SOURCE				
Homo sapiens (human)				
ORGANISM				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE				
1 (bases 1 to 2694)				

AUTHORS	Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE	The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED	14757826
REFERENCE	2 (bases 1 to 2694)
AUTHORS	Mallet,F., Bouton,O. and Oriol,G.
TITLE	Direct Submission
JOURNAL	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France
FEATURES	Location/Qualifiers
source	1..2694 /organism="Homo sapiens" /mol_type="genomic DNA" /isolate="individual 81 allele A" /db_xref="taxon:9606" /chromosome="7" /map="7q21-q22" /sex="female" /cell_type="PBMC's" /note="endogenous_virus: HERV-W" 43..57 /note="splice acceptor site" 284..1900 /note="syncytin" /codon_start=1 /product="envelope glycoprotein" /protein_id="AAQ17591.1" /db_xref="GI:33411045" /translation="MALPYHFLFTVLPSFTLAPPCRCMTSSSPYQBELRWQRFP GNIDAPSVRSLSKGTPTAHTMPCNVCYSATLCMHANTHYWTGKMINPSCPGSLGV TVCTYFTQTGMSGGVQDQAKREKHYKEVISQLTRVHGTSPPYKGLDLSKLHETLR HRLVSLNLTGLHEVSAQNPNCIMCLPNERPVVISIPVQPMNFSFEINTTSV LVGLVSNLEITHTSNLTCVFSNTTYTQDLYSVISKPRKRPVILPFFVIGAGVLGALG YRCNLGSSCFSLFVPPMTIYEQDLYSVISKPRKRPVILPFFVIGAGVLGALG TGIGITTFQFYKLVNSGIVTEKVEIRDIORRAELRNTGPGWLLSOWMPWILP FLGPLAAILLLLPGPCIFNLLVNFVSRIEAVKLQMPKQSKTKIYRRLDRPASP RSDVNDIKGTPEEISAAQPLRPNSAGSS" 1705..1720 /note="splice acceptor site" 1959..2694
misc_feature	
CDS	
ORIGIN	
Query Match	77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity	90.1%; Pred. No. 3e-304;
Matches 1133; Conservative	0; Mismatches 111; Indels 13; Gaps 2;
Qy	1 TCATAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACCTGTTTATTTT 60
Db	
Qy	61 AGGGAAGAATGCTTTAGTATGTTAATCAATCTGGAATCAATTAAGTGAAGAAGTTAAAGA 120
Db	
Qy	121 AATTGAGATCGAATATAATAGTAGAGCAGAGGACCTTCAAAACACTGACCCCTGGGGCT 180
Db	
Qy	1543 AATTCGAGATCGAATACAACTGAGTAGAGAGGAGCTTCGAAACACTGGACCTGGGGCT 1602
Db	
Qy	181 CCTCAGCAATGAGTGCCTGACCTCTCCCTTCTAGGACCTCTAGCAGCTATAATATT 240
Db	
Qy	241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCCTTTGAAGTTTGTCTCTCCAGAAAT 300
Db	
Qy	301 TGAAGCTGAAGCTACAATAATAGTTCTTCAATGGAACCCAGATGCGAGTCAATGACTAA 360
Db	
Qy	1723 CGAAGCTGTAATAACTA-----CAAATGGAGCCCAAGATGCGAGTCCAAGACTAA 1770
Db	
Qy	361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGTCTCTGATGTATGATTAATGACATTGA 420
Db	
Qy	421 AGTCACCCCTCCCGAGGAAATCTCAATGTGCACAAACCCCTACTACACTCTCAATTCAGTAGG 480
Db	
Qy	481 AAGCAGTTAGAGCAGTTGTTCAGCAACCTCCCAACAGTAGTACTTGGGTTTCTCTGTTGAGA 540
Db	
Qy	541 GGGTGGACTGAGACAGAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCNAAAGCCT 600
Db	
Qy	601 ANCTGGGAAGGTGACCCGATCCATCTTTTAAACATGGGGCTTGGAACTTAGCTCACACCCG 660
Db	
Qy	661 ACCAATCAGAGAGCTCACTAAAATGCTAATTCAGGCAAAAACAGGAGGTAAAGCAATAGCC 720
Db	
Qy	721 AATCATCTATTGCTGAGAGCAGACGGGGAAGGACAAAGGATTTGGGATATATAAATCAGGCA 780
Db	
Qy	781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCATCTGCTATGATGGAGCTCTGTTTT 840
Db	
Qy	841 CACTCTATTTCACTCTATTAATCATCACTGACCTCTTCTGCTCGTGTGTTTTTATGG 900
Db	
Qy	901 CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAACAGACCCGCT 960
Db	
Qy	961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGGAGGT 1020
Db	
Qy	1021 ACCCATTTGCCACTCCCGATCAGGCTAAAGGCTTCCCATTTGTTCTGATGCTTAAGTGCC 1080
Db	
Qy	1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTTCCATGGTTCTCTTCCATGAC 1140
Db	
Qy	1141 CCAGCGTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATCTCTTGGTA 1200
Db	
Qy	1201 TCTGTGAGGCAAGAAACCCAGAGTGCAGAGAAAGTGAGGCTTGCCACCATTTGGGAAAG 1257
Db	
Qy	2610 TCCATAAGCCNAGAACCCAGGTCAGAGAAACAGAGGCTTGCCACCATCTTGGGAG 2666
Db	
RESULT 28	
AF52054952	
LOCUS	2694 bp DNA linear PRI 11-FEB-2004
DEFINITION	Homo sapiens individual 95 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION	AF520550
VERSION	AF520550.1
KEYWORDS	GI:33411068
SEGMENT	2 of 2
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.

The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology

Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France

FEATURES

source

1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 95 allele A"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
misc_feature
43..57
/note="splice acceptor site"
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA017597.1"
/gb_xref="GI:33411069"
/translation="MALPYHIFLFTVLLPSFTLTAPPKRCMTSSPYQEFPLWRMQR
TNDAPSYSLSKGTFTTAHPRNCRVHSATLCHNANTHYMTGMINPSCPGGLGV
TWCTVYTGSDGGVODQAREKHVEVLSQTRVHGTSSPYKGLDLSKLHETRT
HRLVSLFNTLTGLHEVSAQNPCTNCWICLPIFRPVYSIPVEOWNFESTINTSV
LVGLVSLNEIHTISNLTCKFSNTITNTNSQCIRWVPTQIVCLPSGIFVFCGTS
YCLNGSSMSFSLFVPPMTIYEDLYSYVSKPRNRPVPLFPVFGVGLGALG
TGIGTITTFQYKLSQBLNGDMERVADSLVLOQLNSLAALVQLNRRALDLTAE
RGCTCLFGEQYVQSGIVTEKVEIKRDIORAEELNTPWMLLSQWMPWILP
FLGPLAAIILLFLGPCIENLLVNFVSSRIEAVKLQMEPKMQSKTIYRPLDRPASP
RSDVNDIKGTPPEISAQPLRPNSAGSS"
misc_feature
1705..1720
/note="splice acceptor site"
1959..2694

CDS

Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
ORIGIN
1 TCAAAATCGAAGAGCTTTAGACTTGCTTAACCGCAAAAGAGGGGGAACCTGTTATTTT 60
1423 TCAAAATCGAAGAGCTTTAGACTTGCTTAACCGCTGAAGAGGGGGAACCTGTTATTTT 1482
61 AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTAATCTGAGAAAGTTAAAGA 120
1483 AGGGGAAGAAATGCTGTTATGTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542
121 AATTGAGATCGAATATAATGTAAGCAGAGGACCTTCAAAACACTGCACCCCTGGGCGCT 180
1543 AATTGAGATCGAATACAACCTGAGCAGAGGAGCTTCGAARCACTGGACCCCTGGGCGCT 1602
181 CCTCAGCAATGAGTGCCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
1603 CCTCAGCAATGAGTGCCTGGATTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1662
241 TTTTACTCTCTTTGGACCCCTGTAATCTTCAACTCTTGTGTTAGTTGTTCTCTCCGAAT 300
1663 GCTACTCTCTTTGGACCCCTGTAATCTTCAACTCTTGTGTTAGTTGTTCTCTCCGAAT 1722

301 TGAAGCTGTAAGCTTACAATAGTCTTTCAATAGAACCCAGAGTGCAGTCCATGACTAA 360
1723 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCAGTCCAGACTAA 1770
361 AATCTACCGTGGAGCCCTTGGACCGGCTGTAGACTATGCTCTGATGTATGATGATGA 420
1771 GATCTACCGCAGACCCCTTGGACCGGCTGTAGCCACGATCTGATGTATGATGATCAA 1830
421 AGTCACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTTACACTCCAATTCAGTAGG 480
1831 AGGACCCCTCTCTGAGGAAATCTAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 1890
481 AAGCAGTTAGAGCAGTGTCTAGCCAACTCTCCCAACAGTACTTGGGTTTTCTGTTGAG 540
1891 AAGCAGTTAGAGCGTCTGCGCCAACTCTCCCAACAGCAGTACTAGGTTTTCTGTTGAG 1950
541 GGTGGAGCTGAGAGCAGGACTAGCTGATTTCTTAGCTGACTAGATCCCNAGCCT 600
1951 TGGGGAGCTGAGAGCAGGACTAGCTGATTTCTTAGCTGACTAGATCCCNAGCCT 2010
601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
2011 AGCTGGGAAGGTGACCATCCCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2070
661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
2071 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
721 AATCATCTATTGCTGAGAGCAGCGGAGGACAGGATTTGGATATAAATACTCAGGCA 780
2131 AATCATCTATTGCTGAGAGCAGCGGAGGACAGGATTTGGATATAAATACTCAGGCA 2190
781 TTCAGCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGTTT 840
2191 TTCAGCAGCGCAACCGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGTTT 2250
841 CACTCTATTTCACCTCTATTAAATCATGCAACTCTTCTGGTCCGTTGTTTATGG 900
2251 CATGCTATTTCACCTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTACGG 2310
901 CTCAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTGTCACCGTCACAGCCCGCT 960
2311 CTTGAGCTGAGCTTTGCTGCGCATCCCACTGCTGTTGTCGCGCACCGAGACCGCC 2370
961 GCTGACTTCCATCCCTTTGGATCCAGAGGTGTCACCTGCTGCTCTGATCCAGCGAGGT 1020
2371 GCTGACTTCCATCCCTTTGGATCCAGAGGTGTCACCTGCTGCTCTGATCCAGCGAGGT 2430
1021 ACCCATTTGCCACTCCCGATCAGGCTAAAGCTTTGCCATTTGCTCTGATGGCTAAGTGCC 1080
2431 ACCCATTTGCCCTCCCAATCGGGCTAAAGGCTTTGCCATTTGCTCTGATGGCTAAGTGCC 2490
1081 TGGTTTTGCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
2491 TGGTTTTGCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550
1141 CCACGCTTCTTAATAGAGCTATAAACAACCTCAACGATGGCCCAAGATTTCCATTTCTTGGTA 1200
2551 CCACGCTTCTTAATAGAGCTATAAACAACCTCAACGATGGCCCAAGATTTCCATTTCTTGGTA 2609
1201 TCTGTGAGGCCCAAGAACCCAGGCTCAGAGAANGTGGGTTGCCACATTTGGGAAG 1257
2610 TCCATTAAGGCCCAAGAACCCAGGCTCAGAGAANGTGGGTTGCCACATTTGGGAAG 2666

RESULT 29

AF520553S2

LOCUS

DEFINITION

AF520554

VERSION

KEYWORDS

AF520553S2 2694 bp DNA linear PRI 11-FEB-2004
Homo sapiens individual 96 allele A, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
AF520554
AF520554.1 GI:33411076

RESULT 31

AF52056352	AF52056352	2694 bp	DNA	linear	PRI 11-FEB-2004
LOCUS	Homo sapiens individual 71 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.				
DEFINITION					
ACCESSION	AF520564				
VERSION	AF520564.1	GI:33411096			
KEYWORDS					
SEGMENT	2 of 2				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 2694)				
	Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.				
TITLE	The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)				
PUBMED	14757826				
REFERENCE	2 (bases 1 to 2694)				
AUTHORS	Mallet, F., Bouton, O. and Oriol, G.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France				
FEATURES	Location/Qualifiers				
source	1..2694				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/isolate="individual 71 allele B"				
	/db_xref="taxon:9606"				
	/chromosome="7"				
	/map="7q21-q22"				
	/sex="male"				
	/cell_type="PBMC's"				
	/note="African				
	endogenous_virus: HERV-W"				
misc_feature	43..57				
	/note="splice acceptor site"				
CDS	284..1900				
	/note="syncytin"				
	/codon_start=1				
	/product="envelope glycoprotein"				
	/protein_id="AAQ17604.1"				
	/db_xref="GI:33411097"				
	/translation="MALPYHFLFTVLLPSFTLAPPRCMTSSSPYQBELWRKQRP GNIDAPSRSLSGTPTPTAHTHPRNCYHSATLCHMANTHWIGKMINPSCPGLGV TVCWYFTQMSDGGVQDQAREKHVEVISQLTRVHGTSSPYKGLDLSKLHETLRT HTRLVSNFTLTGLHEVSAQNPTNCWICLPLNRPYVSIPIVPEQWNFSFEINTTSV LVGPLVSNLEITHTSNLTCVFSNTVTNSQCIRWVTPPTQIVCLPSGIPFVCGTSA YRLNGSSSEMTCSFLVPPMTIITEODLYSVISKPRNKRVPILPFVIGAVGALG TGIGITTSOTFYKLSQELNGDMERVADSLVTIQDOLNSLAVALONRALDLITAE RGGCLFUGESCCYVNSGIVTEKVEIRDRIORRAEELRNTGFWLLSOWMPWILP FLGPLAAIILLLLPFCPIFNLLVNFVFSRIEAVLQMEPRMQSKTIYRRLDRPAS RSDVNDIKGTPEEISAQPLLRPNASGS"				
misc_feature	1705..1720				
	/note="splice acceptor site"				
LTR	1959..2694				
ORIGIN					
Query Match	77.9%; Score 1035.2; DB 9; Length 2694;				
Best Local Similarity	90.1%; Pred. No. 3e-304;				
Matches 1133; Conservative	0; Mismatches 111; Indels 13; Gaps 2;				
Qy	1 TCAAAATCGAAGCTTTAGACTTGCTAAACCGCCAAAGAGAGGGGAACTGTTATTTT 60				
Db	1423 TCAAAATCGAAGCTTTAGACTTGCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 1482				
Qy	61 AGGGGAAGAAGCTCTTGTAGTATGTTAATCAATCTGGAATCATTAAGAGAAAGTTAAAGA 120				
Db	1483 AGGGGAAGAAGCTCTTGTATGTTAATCAATCCGAATCGTCACTGAGAAGTTAAGA 1542				
Qy	121 AATTGAGATCGAATATTAATGTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGCGCT 180				

```
Db      2610 TCCATAAGGCCAAGAACCCAGGTCAGAGAACACAGGAGCTTGCCACCATCTTTGGGAG 2666

RESULT 32
AX329572
LOCUS      56093 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION      Sequence 81 from Patent WO0194629.
ACCESSION      AX329572
VERSION        AX329572.1  GI:18102550
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1. Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL      Patent: WO 0194629-A 81 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source      1..56093
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      77.9%; Score 1035.2; DB 6; Length 56093;
Best Local Similarity 90.1%; Pred. No. 3.2e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;

QY      1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTATTTT 60
Db      37018 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGGAACCTGTTATTTT 37077

QY      61 AGGGGAAGAGTCTGTTAGTATCTTAATCAATCTGGAATCATTAATCTGAGAAAGTTAAAGA 120
Db      37078 AGGGGAAGAGTCTGTTATTTATGTTTAAATCAATCCGAAATGCTGACAGAAAGTTAAAGA 37137

QY      121 AATTGAGATCGAATATAATGTAGACGAGGAGACCTTCAAAACACATGCACCCCTGGGGCCT 180
Db      37138 AATTGAGATCGAATACAACTAGACGAGGAGGCTTCAAAACACATGCACCCCTGGGGCCT 37197

QY      181 CCTCAGCAATGATGCGCTGGAGCTCTCCCTTTCTTAGGACCTCTAGGACCTATATATTT 240
Db      37198 CCTCAGCAATGATGCGCTGGAGCTCTCCCTTTCTTAGGACCTCTAGGACCTATATATTT 37257

QY      241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 300
Db      37258 GCTACTCTCTTTGGACCCCTGTATCTTCAACTTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 37317

QY      301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCGATGCCATGACTAA 360
Db      37318 CGAAGCTGTAAAGCTA-----CAATGGAGCCCAAGATGCGATGCCATGACTAA 37365

QY      361 AATCTACCTGGAGCCCTGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db      37366 GATCTACCGGAGACCCCTGGACCGGCTGTAGCCCAAGATGCGATGTTAATGACATGA 37425

QY      421 AGTCACCCCTCCGAGGAAATCTCACTGCACAAACCCCTACTACATCCCAATTCAGTAGG 480
Db      37426 AGGCACCCCTCCGAGGAAATCTCAGTGCACAAACCCCTACTACTAGCCCAATTCAGCAGG 37485

QY      481 AAGCAGTTAGAGGAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
Db      37486 AAGCAGTTAGAGGAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 37545

QY      541 GGGTGGACTGAGAGAGGAGTACTAGTGGATTTCTTAGGCTGACTAAGAAATCCCNAGCCT 600
Db      37546 TGGGGGACTGAGAGAGGAGTACTAGTGGATTTCTTAGGCTGACTAAGAAATCCCNAGCCT 37605

QY      601 ANCTGGGAAGGTGACCCGCATCCATCTTTTAAACATGGGGCTTGGCAACTTAGCTCACACCCG 660

Db      37606 AGCTGGGAAGGTGACCCGCATCCATCTTTAAACACGGGGCTTGGCAACTTAGCTCACACCTG 37655

QY      661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAGAGGTAAGAGCAATAGCC 720
Db      37666 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAGAGGTAAGAGCAATAGCC 37725

QY      721 AATCATCTATTGCTGAGAGCAGCGGGAAGGAGGATTTGGATATATAAATCAGGCA 780
Db      37726 AATCATCTATTGCTGAGAGCAGCGGGAAGGAGGATTTGGATATATAAATCAGGCA 37785

QY      781 TTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTT 840
Db      37786 TTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTT 37845

QY      841 CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTGTTTATGG 900
Db      37846 CATGCTATTTCATCTATTAAATCTTGCACTGCACTCTTCTGGTCCATGTTTCTTACGG 37905

QY      901 CTCGAAGTGAGCTTTTGTTCGCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCT 960
Db      37906 CTTGAGCTGAGCTTTTCGCTCGCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCT 37965

QY      961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGCTCTGATCCAGCAGCT 1020
Db      37966 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGCTCTGATCCAGCAGGCT 38025

QY      1021 ACCCAATTCGCACTCCCGATCAGGCTAAAGGCTTGGCAATTTCTTCGATGGCTTAAGTGCC 1080
Db      38026 ACCCAATTCGCGCTCCCAATCGGCTTAAGGCTTGGCAATTTCTTCGATGGCTTAAGTGCC 38085

QY      1081 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTCACTGGGTTCCATGTTCTTCCATGAC 1140
Db      38086 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTCACTGGGTTCCATGTTCTTCTCTGAC 38145

QY      1141 CCACGGCTTCTTAATAGACTATAACACTCACCAGATGCGCCGAGATTTCCATCTCTTGGTA 1200
Db      38146 CCACAGCTTCTTAATAGACTATAACACTCACCAGATGCGCCGAGATTTCCATCTCTTGGTA 38204

QY      1201 TCTGTGAGGCCAAGAACCCAGCTCAGAGAANGTGAGGCTTGCCACCATTTGGGAAG 1257
Db      38205 TCCATAAGGCCAAGAACCCAGCTCAGAGAANGTGAGGCTTGCCACCATTTGGGAG 38261

RESULT 33
HSAC000064      56093 bp      DNA      linear      PRI 13-NOV-1996
LOCUS      Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
DEFINITION
ACCESSION      AC000064
VERSION        AC000064.1  GI:1669369
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1. Pauley, A.
The sequence of H. sapiens BAC clone RG083M05
Unpublished (1996)
REFERENCE
2. (bases 1 to 56093)
Waterson, R.
Direct Submission
Submitted (13-NOV-1996)
JOURNAL      Genome Sequencing Center
COMMENT      Department of Genetics, Washington University
St. Louis, MO 63108, USA
e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
sections once, or longer because we provide a small overlap between
neighboring submissions.
```


This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:

This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
VECTOR: pBEO
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H_RG083M05; actual end is at 56093 of H_RG083M05

This clone contains STS sWSS1725.

FEATURES

source

Location/Qualifiers

1..56093

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="7"

/map="7q21.7q22"

/clone="H_RG083M05"

/clone.lib="CITB-978SK-B"

complement(838..1131)

/rpt_family="ALU"

<1360..16971

/gene="WUGSC:H_RG083M05.1"

join(<1360..1503,4181..4370,4587..4774,6422..6556,

9483..9547,11631..11773,11864..12021,13131..13296,

14885..14988,16349..16546,16837..16971)

/gene="WUGSC:H_RG083M05.1"

/note="Atpase; strong similarity to peroxisome biosynthesis protein PAB1 (PID:g1172019); coded for by human cDNA C04279 (NID:g1467530)"

/codon_start=1

/protein_id="BAB46346.1"

/db_xref="GI:1669371"

/translation="KRLNIQKTEVASEAVWQPSVVLDDLDLGLPAVPEHEH

SPDAQRCEILCNVTKKDCDINKFTDLDQHVAKETGGFVARDFTVLVDRLHSRL

SQSISTRKLVLTLDQKALRFLPASRLVNLKPRDLGWDKIGLHVEVRQILMD

TIQPAKCLKKERYPELFANLPQRQTGILLYGPGTGTLLAGVIAREGRMFVSV

KGPILLSKYIGASEQVRDIETIRQAAPKPCILFPDEPSIAPRRGHNTGVTDRVNVQ

LLTOLDGVEGLOGVYVLAATSRPDLIDPALLRPGRLDKVCVCPDPDQVTIISLSEKTC

QMLRSFLVSRLEILNVLSLPLADDVDLQHVASVTDSTGTADLKALLYNQALELHG

MLLSKMSILPDESFPNRYLFSSYSELSNGTSSQCLASPSMTQDLPQVPI

GKQQLFSPQVPLRTASQGCQELTQEQRDQLRADISIITKGYRSQSGEDEMNPQGP

KTRLAISQHLMTALGHTRPISIEDDKNPAEL"

complement(4948..5130)

/rpt_family="ALU"

complement(6581..7133)

/rpt_family="L1"

complement(7767..8037)

/rpt_family="ALU"

complement(8186..8472)

/rpt_family="ALU"

8473..8625

/gene="WUGSC:H_RG083M05.1"

/note="match to human 3' EST H75782 (NID:g1049794), bases

287-444"

8841..9161

/gene="WUGSC:H_RG083M05.1"

/note="match to human 5' EST H75921 (NID:g1050050), bases

21-348"

9481..9547

/gene="WUGSC:H_RG083M05.1"

/note="match to human 5' EST N22627 (NID:g1130501), bases

276-343"

complement(12612..12907)

/rpt_family="ALU"

13670..13793

/gene="WUGSC:H_RG083M05.1"

/note="match to human 5' EST H41382 (NID:g117434), bases

143-266"

13794..13877

/rpt_family="ALU"

13878..13906

/gene="WUGSC:H_RG083M05.1"

/note="match to human 5' EST H41382 (NID:g117434), bases

30-58"

13907..14104

/rpt_family="ALU"

complement(14110..14137)

/rpt_family="L1"

complement(15618..15907)

/rpt_family="ALU"

17227..17522

/rpt_family="ALU"

18667..19235

/note="match to human fetal brain 5' EST D61494

(NID:g970409), bases 1-255, and to human 3' EST R07476

(NID:g759399)"

19550..19670

/rpt_family="ALU"

21507..37303

/note="similarity to various SS-RNA virus polyproteins;

pseudogene; region of matches and close matches to

multiple human ESTs, see R68740 (NID:g842257)"

37316..37489

/note="Grail prediction, score = 80"

/evidence=not_experimental

complement(38938..39224)

/rpt_family="ALU"

39225..39707

/note="match to multiple human ESTs, see N30113

(NID:g1148633)"

39800..40085

/rpt_family="ALU"

complement(40247..40538)

/rpt_family="ALU"

complement(40632..40924)

/rpt_family="ALU"

complement(42283..42891)

/rpt_family="ALU"

complement(45474..45613)

/rpt_family="ALU"

complement(45614..45737)

/note="match to human 3' EST H48898 (NID:g988738), bases

129-333"

complement(46107..47026)

/note="match to multiple human ESTs, see N81064

(NID:g1243765), H48897 (NID:g988737), and W78831

(NID:g273146)"

complement(47027..47318)

/rpt_family="ALU"

complement(47365..47782)

/note="match to multiple human ESTs, see W37495

(NID:g1319089)"

47898..48115

/note="match to human 5' EST H62306 (NID:g1015138), bases

93-368"

complement(48116..48405)

/rpt_family="ALU"

complement(48406..48584)

/note="match to human 3' EST N29952 (NID:g1148472), bases

290-455, and 5' EST R12730 (NID:g765806)"

complement(48787..49405)

/rpt_family="ALU"

complement(49406..49534)

/note="match to human 3' EST R65794 (NID:g838432), bases

Db	37606	AGCTGGGAAGTGAACAATCACTTAAACACGGGGCTTGAACCTTAGCTCACACCTG	37666
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACACGAGGCTAAAGCAATAGCC	720
Db	37666	ACCAATCAGAGAGCTCACTAAATGCTTAATTAGGCAAAACACGAGGCTAAAGCAATAGCC	37725
Qy	721	AATCATCTATTCGCTGAGAGCACGGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780
Db	37726	AATCATCTATTCGCTGAGAGCACGGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	37785
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCCCTCCCATTTGTATGGAGCTCTGTTTT	840
Db	37786	TTGAGCGCGCAACGGCAACCCCTTTGGGTCCCCCTCCCATTTGTATGGAGCTCTGTTTT	37845
Qy	841	CACCTATTTTCACTCTATTAAATCATGNAATGCACTCTCTCTGTCCTGTTTTTATGG	900
Db	37846	CATGCTATTTCACTCTATTAAATCTTGCAACTGCACCTCTCTGTCCTCAATGTTCTTACGG	37905
Qy	901	CTCAAGCTGAGCTTTTGTTCGGCATCCACCACTGCTGTTTGGCCACCGTCAGACCCGCT	960
Db	37906	CTTGAGCTGAGCTTTTGTTCGGCATCCACCACTGCTGTTTGGCCACCGTCAGACCCGCT	37965
Qy	961	GCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCGAGT	1020
Db	37966	GCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCGAGT	38025
Qy	1021	ACCCATTTGCCATCTCCGATCAGGCTTAAGGCTTGGCATTTGTCCTGCAATGGCTAAGTCC	1080
Db	38026	ACCCATTTGCCATCTCCGATCAGGCTTAAGGCTTGGCATTTGTCCTGCAATGGCTAAGTCC	38085
Qy	1081	TGGGTTTGCTTAATAGAACTGAACACTGGTCACTGGGTTCATGGTCTCTTCCATGAC	1140
Db	38086	TGGGTTTGCTTAATAGAACTGAACACTGGTCACTGGGTTCATGGTCTCTTCCATGAC	38145
Qy	1141	CCAGCGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA	1200
Db	38146	CCAGCGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA	38204
Qy	1201	TCTGTGAGGCCAAGAACCCCGAGTTCAGAGAANTGAGGCTTGCCACCATTTGGGAAG	1257
Db	38205	TCCATAGGCCAAGAACCCCGAGTTCAGAGAANTGAGGCTTGCCACCATTTGGGAAG	38261
RESULT 34			
AC007566/c			
LOCUS	AC007566	149194 bp	DNA linear PRI 01-MAR-2002
DEFINITION	Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.		
ACCESSION	AC007566		
VERSION	AC007566.2	GI:11181861	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 149194)		
TITLE	Sulston,J.E. and Waterston,R.		
JOURNAL	Toward a complete human genome sequence		
MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)		
PUBMED	99063792		
REFERENCE	9847074		
AUTHORS	2 (bases 1 to 149194)		
TITLE	Du,Z.		
JOURNAL	The sequence of Homo sapiens BAC clone CTB-10G5		
REFERENCE	Unpublished (2001)		
AUTHORS	3 (bases 1 to 149194)		
TITLE	Waterston,R.H.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (15-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
AUTHORS	4 (bases 1 to 149194)		
TITLE	Waterston,R.		

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (02-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 149194)
Waterston,R.H.
TITLE
JOURNAL

Direct Submission
Submitted (16-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 149194)
Waterston,R.H.
TITLE
JOURNAL

Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
7 (bases 1 to 149194)
Waterston,R.H.
TITLE
JOURNAL

Direct Submission
Submitted (06-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
8 (bases 1 to 149194)
Waterston,R.
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics

Center project name: H_RG010G05

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
Clone CTB-10G5 is from the first release of the human BAC library CITB-97ASK-B. The library contains cloned DNA from the male fibroblast cell line 97ASK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
VECTOR: pBelOBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

FEATURES
source
The clone CTB-10G5 contains the entire sequence of CTB-83M5.
1..149194
Location/Qualifiers
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/clone="CTB-10G5"
/clone_lib="CITB-97ASK-B"
1..1634
/rpt_family="L1"
repeat_region
2248..2388
misc_feature
/notes="match to EST BG752883 (NID:g14063536)"
misc_feature
/notes="match to EST AA149693 (NID:g1720635) zn99d06.rl"
misc_feature
/notes="match to EST AW579261 (NID:g7254310)"
misc_feature
/notes="match to EST BG766882 (NID:g14077535)"
misc_feature
/notes="match to EST BI160365 (NID:g14620366)"
2248..2374
misc_feature
/notes="similar to Homo sapiens EST BF758865 (NID:g12106765)"
misc_feature
/notes="match to EST AV686676 (NID:g10288539)"
misc_feature
/notes="match to EST AU123510 (NID:g10948226)"
misc_feature
/notes="match to EST AV686676 (NID:g10288539)"
repeat_region
2696..3066
/rpt_family="MaLR"
repeat_region
3108..3332
/rpt_family="Alu"
3540..3628
misc_feature
/notes="match to EST AA425526 (NID:g2106267) zw48b03.rl"
misc_feature
/notes="similar to Mus musculus EST BB253526 (NID:g8946272)"
misc_feature
/notes="match to EST AV686676 (NID:g10288539)"
misc_feature
/notes="match to EST AA149693 (NID:g1720635) zn99d06.rl"
misc_feature
/notes="match to EST AU123510 (NID:g10948226)"
misc_feature
/notes="match to EST BG766882 (NID:g14077535)"
misc_feature
/notes="match to EST BI160365 (NID:g14620366)"
misc_feature
/notes="match to EST AW579261 (NID:g7254310)"
misc_feature
/notes="match to EST BE272564 (NID:g9146913)"
misc_feature
/notes="match to EST BE299708 (NID:g9183456)"
3714..3785
misc_feature
/notes="match to EST BG260659 (NID:g12770475)"
3717..3785
misc_feature
/notes="similar to Mus musculus EST BE994936 (NID:g10678674)"
3719..3785
misc_feature
/notes="match to EST AA149693 (NID:g1720635) zn99d06.rl"
3719..3785
misc_feature
/notes="match to EST AA425526 (NID:g2106267) zw48b03.rl"
3719..3785
misc_feature
/notes="match to EST AU123510 (NID:g10948226)"
3719..3785
misc_feature
/notes="match to EST BE272564 (NID:g9146913)"
3719..3785
misc_feature
/notes="match to EST BE299708 (NID:g9183456)"
3719..3785
misc_feature

		/note="match to EST BG766882 (NID:g14077535)"	
misc_feature	3719..3785	/note="match to EST B1160365 (NID:g14620366)"	
misc_feature	3719..3785	/note="similar to Mus musculus EST BB253526 (NID:g8946272)"	
misc_feature	3860..3862	/note="match to EST AV686676 (NID:g10288539)"	
repeat_region	4730..4827	/rpt_family="L2"	
repeat_region	4834..4925	/rpt_family="L1"	
misc_feature	5715..5860	/note="match to EST AA425526 (NID:g2106267) zw48b03.z1"	
misc_feature	5715..5860	/note="match to EST AU123510 (NID:g10948226)"	
misc_feature	5715..5860	/note="match to EST BE272564 (NID:g9146913)"	
misc_feature	5715..5860	/note="match to EST BE299708 (NID:g9183456)"	
misc_feature	5715..5860	/note="match to EST BG260659 (NID:g12770475)"	
misc_feature	5715..5860	/note="match to EST BG766882 (NID:g14077535)"	
misc_feature	5715..5860	/note="match to EST B1160365 (NID:g14620366)"	
misc_feature	5715..5860		
Query Match	77.9%;	Score 1035.2; DB 9; Length 149194;	
Best Local Similarity	90.1%;	Pred. No. 3.2e-304;	
Matches 1133; Conservative	0; Mismatches 111; Indels 13; Gaps 2;		
QY	1	TCAAAATCGAAGAGCTTTAGACTTCTTAACGCGCCAAAAGAGGGGAACCTGTATTATTTT	60
Db	84986	TCAAATTCGAAGAGCTTTAGACTTCTTAACGCGCTGNAAGAGGGGAACCTGTATTATTTT	84927
QY	61	AGGGGAAGAATCCTGTAGTAGTTAAATCAATCTGGAATCATTACTGAGAAAGTTAAGA	120
Db	84926	AGGGGAAGAATCCTGTATTATGTTAAATCAATCCGAAATCGTCACTGAGAAAGTTAAGA	84867
QY	121	AATTTGAGATCGAATATAATGTAGACGACGAGCACCTTCAAACACTGCACCCTCGGGCCT	180
Db	84866	AATTCGAGATCGAATACAACGCTAGACGACGAGAGCTTGAAACACTGGAACCTCGGGCCT	84807
QY	181	CCTCAGCCAATGGATGCCCTGGACTCTCCCCTTCTTAGACCTCTAGCAGCTATAATATT	240
Db	84806	CCTCAGCCAATGGATGCCCTGGATTCTCCCCTTCTTAGACCTCTAGCAGCTATAATATT	84747
QY	241	TTTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTGTTAAGTTTGCTCTTCCAGAAAT	300
Db	84746	GCTACTCTCTTTGGACCCCTGATCTTTAAACCTCTTGTAACTTGTCTCTTCCAGAAAT	84687
QY	301	TGAAGCTGTAAAGCTTACAATAAGTTCTTCAAAATGAAACCCAGATGCAATCCATGACTAA	360
Db	84686	CGAAGCTGTAAAACTA-----CAAAATGGAGCCCAAGATGCAATCCAGACTAA	84639
QY	361	AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGCAATTGA	420
Db	84638	GATCTACCGACAGCCCTGGACCGGCTGCTAGCCCAAGATCTGATGTTAATGCAATCAA	84579
QY	421	AGTCACCCCTCCCGAGGAAATCTCAACTGSCAACCCCTTACTACTCAATTCAGTAGG	480
Db	84578	AGGCACCCCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG	84519
QY	481	AAGCAGTTAGACAGTTGTCAGCCAACTTCCCCAAGACAGTACTTGGTTTTCTGTTTGA	540
Db	84518	AAGCAGTTAGACGGTCTGTCGGCAACCTTCCCAACAGCACTTAGGTTTTCTGTTTGA	84459
QY	541	GGGTGACCTGAGACAGACAGTACTGTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT	600
Db	84458	TGGGGACTCTGAGAGA CAGACTAGCTGTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCT	84399
QY	601	ANCTGGGAAGGTGACCGGCATCCATCTTTAAACATCGGGGCTTCGAACTTAGCTCACACCGC	660

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

FEATURES	d'Italie, Lyon 69364 cedex 07, France	
	Location/Qualifiers	
source	1..2694	
	/organism="Homo sapiens"	
misc_feature	/mol_type="genomic DNA"	
	/isolate="individual 22 allele B"	
CDS	/db_xref="taxon:9606"	
	/chromosomes="7"	
misc_feature	/map="7q21-q22"	
	/sex="male"	
CDS	/cell_type="PBMC's"	
	/note="Caucasian"	
misc_feature	endogenous_virus: HERV-W"	
	43..57	
CDS	/note="splice acceptor site"	
	284..1900	
misc_feature	/note="syncytin"	
	/codon_start=1	
CDS	/product="envelope glycoprotein"	
	/protein_id="AAQ17568.1"	
misc_feature	/db_xref="GI:33410953"	
	/translation="WALPYHFLFTVLLPSFTLTAPPCCRCMTSSPYQVFLWRQRP"	
CDS	GNIDPSYSLSGKGTPTFTAHPRNCVHSATLCMHANTHYWTGMINPSPCGGLGV	
	TVCWTYFTQSGDGGVODQAREKHVEVISQLTOVHGTSPTKGLDLKSLHETLR	
misc_feature	HRLVSLFTNLTGLHEVSAQNPNCWICLPNFRVYSI PVPEQWNESTREINTSV	
	LVGLVSLNLEIHTHSLNLCVKSNTYITNSQCIKRWPTPTQIVCLPSGIFPVCGTSA	
misc_feature	YCLNGSSMCFSLSPVPMPTIYEDQLYNYSIKPRNKRVPILPFPVIGAGVIGALG	
	TGIGITTSQYKLSBELNGDMERVADSLVTLQDLSLAALVONRRALDLTLAE	
misc_feature	RGCTCLFGECCYVYVNSQIVTEKVEIRIQRAEELRNTGPMGLLSQWMPWILP	
	FLGLAAIITLLLFPGICFNLLVNFVSSRIEAVKLQMEPKMQSKTIYRRPLDRASP	
misc_feature	RSDVNDIKGTPPEEISAAQPLLRPNSSGS"	
	1705..1720	
misc_feature	/note="splice acceptor site"	
	1959..2694	
LTR	ORIGIN	
	Query Match 77.8%; Score 1033.6; DB 9; Length 2694;	
LTR	Best Local Similarity 90.1%; Pred. No. 9.2e-304;	
	Matches 1132; Conservative 0; Mismatches 112; Indels 13; Gaps 2;	
QY	1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTTATTTT 60	
	1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGGAACCTGTTTATTTT 1482	
QY	61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA 120	
	1483 AGGGGAAGAAATGCTGTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542	
QY	121 AATTGAGATCGAATATATGATAGCAGAGGACCTTCAAAACACTGCACCCCTGGGGCCT 180	
	1543 AATTGAGATCGAATACAACTAGCAGAGGAGCTTCGAACACTGGACCCCTGGGGCCT 1602	
QY	181 CCTCAGCAATGATGCTGCTGCTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATATT 240	
	1603 CCTCAGCAATGATGCTGCTGCTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATATT 1662	
QY	241 TTTACTCTCTTTGGACCCCTGATPCTTCAACTCTCTGTTAAAGTTGCTCTCTCCAGAA 300	
	1663 GCTACTCTCTTTGGACCCCTGATPCTTCAACTCTCTGTTAAAGTTGCTCTCTCCAGAA 1722	
QY	301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCGATCCATGACTAA 360	
	1723 CGAAGCTGTAAACTA-----CAAAATGGAACCCAGATGCGATCCATGACTAA 1770	
QY	361 AATCTACCTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420	
	1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830	
QY	421 AGTCACCCCTCCGAGAGAAATCTCAACTGCACAAACCCCTACTACATCCCAATTCAGTAG 480	
	1831 AGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890	
QY	481 AAGCAGTTAGGAGCTTGTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540	
Db	1891 AAGCAGTTAGGAGCTTGTGTCAGCAACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGA 1950	
QY	541 GGGTGGACTGAGAGCAGAGCTAGCTGATTTCTTAGGCTGACTAAGAAATCCCNAGCCCT 600	
	1951 TGGGGGACTGAGAGCAGAGCTAGCTGATTTCTTAGGCTGACTAAGAAATCCCNAGCCCT 2010	
QY	601 ANCTGGGAAGGTGACCGATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACAACCG 660	
	2011 AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACAACCTG 2070	
QY	661 ACCAATCAGAGAGCTCCTATAAATGCTAATCAGGCAAAACAGGAGGTTAAGCAATAGCC 720	
	2071 ACCAATCAGAGAGCTCCTATAAATGCTAATCAGGCAAAACAGGAGGTTAAGCAATAGCC 2130	
QY	721 AATCATCTATTGCTGAGAGCAGCGGAGGAGCAAGGATGGGATATAAATCTCAGGCA 780	
	2131 AATCATCTATTGCTGAGAGCAGCGGAGGAGCAAGGATGGGATATAAATCTCAGGCA 2190	
QY	781 TTCAAGCCAGCAACCAACCCCTTTGGGTCCCTCCCATTTGTATGGGAGCTCTGTTTT 840	
	2191 TTGAGCCGCGCAACCGCAACCCCTTTGGGTCCCTCCCATTTGTATGGGAGCTCTGTTTT 2250	
QY	841 CACTCTATTCTCACTCTATTAAATCATGCAATGCACTCTTCTGGTCCGTTTATGG 900	
	2251 CATGCTATTCTCACTCTATTAAATCATGCACTCTTCTGGTCCGTTTATGG 2310	
QY	901 CTCAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGTCACAGCCGCT 960	
	2311 CTTGAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGTCACAGCCGCT 2370	
QY	961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCAGGT 1020	
	2371 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCAGGT 2430	
QY	1021 ACCAATGTCCTCCCGATCAGGCTAAAGGCTTGCATTTGCTCTGATGCTCTGATGCTCC 1080	
	2431 ACCAATGTCCTCCCGATCAGGCTAAAGGCTTGCATTTGCTCTGATGCTCTGATGCTCC 2490	
QY	1081 TGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140	
	2491 TGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550	
QY	1141 CAAGGCTTCTTAATAGAGCTATAAACACTCAGCAGTGGCCCAAGATTCCATTTCTTGGTA 1200	
	2551 CCACGCTTCTTAATAGAGCTATAAACACTCAGCAGTGGCCCAAGATTCCATTTCTTGGTA 2609	
QY	1201 TCTGTGAGCCCAAGAACCCCGATCAGAAAGTGAAGGCTTGCACCATTTGGGAAG 1257	
	2610 TCCATAAGCCCAAGAACCCCGATCAGAAAGTGAAGGCTTGCACCATTTGGGAAG 2666	
AF520499S2	2694 bp DNA linear PRI 11-FEB-2004	
	Homo sapiens individual 30 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.	
AF520500	AF520500.1 GI:33410968	
	2 of 2	
Homo sapiens (human)	Homo sapiens	
	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 2694)	
REFERENCE	Mallat, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L., and Mandrand, B.	
	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology	
AUTHORS	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
	14757826	
TITLE	2 (bases 1 to 2694)	
JOURNAL		
PUBMED		
REFERENCE		

Fr1 Feb 25 16:26:27 2005

AUTHORS	Mallet, F., Bouton, O. and Oriol, G.	
	Direct Submission	
TITLE	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142	
	CNRS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée	
JOURNAL	d'Italie, Lyon 69364 cedex 07, France	
	Location/Qualifiers	
FEATURES	1..2694	
	/organism="Homo sapiens"	
source	/mol_type="genomic DNA"	
	/isolate="individual 30 allele B"	
	/db_xref="taxon:9606"	
	/chromosome="7"	
	/map="7q21-q22"	
	/sex="female"	
	/cell_type="PBMC's"	
	/note="Caucasian"	
misc_feature	endogenous_virus: HERV-W"	
	43..57	
	/notes="splice acceptor site"	
	284..1900	
CDS	/note="syncytin"	
	/codon_start=1	
	/product="envelope glycoprotein"	
	/protein_id="AAQ17572.1"	
	/db_xref="GI:33410969"	
	/translation="WALPYHIFLFTVLPSFTLTAPPCCRCMTSSSPYQEFLLWRMQR GNIDAPYSRLSKGPTFTFAHTRPNRCVHSATLQWANTHYWTKGLDLSKLHETLR TWCTWYPTGSDGGVQDOAREKHVKEVISQLTRVHTGTSPPYKGLDLSKLHETLR HTRVLSFNTITGLHEVSAQPTNCWICLPLNRPYVSIPIVPEQNNFSTEINTSV LVGPLVSNLEITCSNLCTCKFSNTIITNSOICIKWVPTQIVCLPSGIFPVGCTSA YVCLNGSSSMCFSLSPFPMPTIYEQDLSVISKPRNRPVILPFIWAGVLGALG TGIGTITSTQFYKLSQELNGMEVADSLVTLQDLSLAVALVQNRRLDLTLTA RGCTGLFGECCYVYNQSGIVTEKVKIEIRQRAEELRTGFWLLSQMWPILP FLGPLAAIILLILFGPCIENLLVNFVSSRIEAVKLQMBPKMQSKTKYIRPLDRPASP RSPVDIKCTPEPEISAAQPLLRPNISAGSS" 1705..1720 /note="splice acceptor site"	
	1959..2694	
LTR	/notes="splice acceptor site"	
	1959..2694	
ORIGIN	77.8%; Score 1033.6; DB 9; Length 2694;	
	Query Match 90.1%; Pred. No. 9.2e-304; Indels 13; Gaps 2;	
Matches 1132; Conservative 0; Mismatches 112; Indels 13; Gaps 2;		
1	TCAAAATCGAAGAGCTTTAGACTTGCTAAACCCGCAAGAGGGGGAACTGTTTATTTT 60	
	1423 TCAAAATCGAAGAGCTTTAGACTTGCTAAACCCGCTGAAGAGGGGGAACTGTTTATTTT 1482	
61	AGGGGAGAGTCTGTTAGTATCTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120	
	1483 AGGGGAGAGTCTGTTAGTATCTTAATCAATCTGGAATCTGACATGAGAAAGTTAAAGA 1542	
121	AATTTCAGATCGAATAATAATGTAGAGCAGAGGACCTTCAAAACACCTGCACCCCTGGGGCT 180	
	1543 AATTTCAGATCGAATAACAAGTAGAGCAGAGGAGCTTGGAAACACCTGGACCCCTGGGGCT 1602	
181	CCTCAGCAATGAGATGCCCTGGACTCTCCCTCTTAGACCTCTAGCAGCTATATATT 240	
	1603 CCTCAGCAATGAGATGCCCTGGACTCTCCCTCTTAGACCTCTAGCAGCTATATATT 1662	
241	TTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTTTAAAGTTTGTCTCTTCCAGAAAT 300	
	1663 GCTACTCTCTTTGGACCTGTATCTTAACTCTCTTTTAACTTTGTCTCTTCCAGAAAT 1722	
301	TGAAGCTGTAAAGCTACAAATAGTCTTTTAAATGGAAACCCAGATGAGTCCATGACTAA 360	
	1723 CGAAGCTGTAAAGCTA-----CAAATGGAGCCCAAGATGCAGTCCAGACTAA 1770	
361	AATCTACCGTGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAATGACATTTGA 420	
	1771 GATCTACCGGACCCCTGGACCGGCTCTAGCCACGATCTGATGTTAATGACATCAA 1830	
421	AGTACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTTCCAATTCAGTAGG 480	

in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
2 (bases 1 to 2694)
Mallet, F., Bouton, O. and Oriol, G.
Direct Submission
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolates="individual 34 allele A"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
misc_feature
43..57
/notes="splice acceptor site"
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA017573.1"
/db_xref="GI:33410973"
/translation="MALPYHIFLTVLLPSFTLTAPPCCRMWSSPSYQEFLLWRMORP
GNIDAPYRSLSKGTFTFTAHMPCNVHSATLCMHANTHYMTGMINPSPGGLGV
TWCVTFQTTGSDGGVODQAREKHVEVISOLTRVHGTSPYKGLDLKSLHETLR
HPLVSLFNTLTGLHEVSAONPTNCWICLPLNFRPVSVIPVEOWNFSTELNITSV
LVGLVSNLEIHTISNLTCKFSNTYTTNSQCIRWVTPPTQIVCLPSLFFVCGTSA
YCLNGSSMCMFSLFVLPWMTIYEQDLYSVISPRNKRVPILFVIGAGVGLAG
TGIGTITTSQFYKLSBELNGMERVADSLVLOQLNSLAALVQLNRALDLTAE
RGCTCLFLGCEGVYVNGSIVTEKVEIKRDRIRAEELRNTGPMGLLSQWMPWLP
FLGLAAIILLGLFGPCIFNLVNFVSRIEAVKLQWEPKMQSKTIYRPLDRPASP
RSDVNDIKGTFPEISAAQPLRPNSAGSS"
misc_feature
1705..1720
/notes="splice acceptor site"
1959..2694
LTR
ORIGIN
Query Match 77.8%; Score 1033.6; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 9.2e-304;
Matches 1132; Conservative 0; Mismatches 112; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTTATTTT 60
DB 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGAACTGTTATTTT 1482
QY 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCATTAAGAGAAAGTTAAAGA 120
DB 1483 AGGGGAAGATGCTGTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542
QY 121 AATTGAGATCGAATATAAGTAGAGAGAGGACCTTCAAAACACTGCACCCCTGGGCGCT 180
DB 1543 AATTGAGATCGAATATAAGTAGAGAGAGGAGCTTCGAAACACTGCACCCCTGGGCGCT 1602
QY 181 CCTCAGCAATGATGCTGCTGCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
DB 1603 CCTCAGCAATGATGCTGCTGCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 1662
QY 241 TTTACTCTCTTTGGACCCGTATCTTCAACTCTCTTGTGTTAAGTTGCTCTTCCGAAT 300
DB 1663 GCTACTCTCTTTGGACCCGTATCTTCAACTCTCTTGTGTTAAGTTGCTCTTCCGAAT 1722
QY 301 TGAAGCTGTAAGCTACAATAGTCTTCAATGGAACCCAGATGCGAGTCCATGACTAA 360
DB 1723 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCGAGTCCGAAGACTAA 1770
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATCTTAATGACATTGA 420

DB 1771 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTAATGACATCA 1830
QY 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCAACACCCCTACTACTCACTCAATTCAGTAG 480
DB 1831 AGGCACCCCTCTCTGAGGAAATCTAGCTGCAACACCTCTACTACGCCCAATTCAGCAG 1890
QY 481 AAGCAGTTAGAGCAGTTGTACGCCAACCTCCCAACAGTACTTGGTGTCTCTCTGTTGAGA 540
DB 1891 AAGCAGTTAGAGCGGTGCTCGGCCAACCTCCCAACAGCAGTACTTGGTGTCTCTCTGTTGAGA 1950
QY 541 GGTGTGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCCT 600
DB 1951 TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCTAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCTTTAAACATGCGGCTTGCAACTTAGCTACACCCG 660
DB 2011 AGCTGGGAAGGTGACCATCTTTAAACACGCGGCTTGCAACTTAGCTACACCTG 2070
QY 661 ACCAATCAGAGAGCTCACTAAATGCTATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
DB 2071 ACCAATCAGAGAGCTCACTAAATGCTATTAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
QY 721 AATCATCTATTCCTGAGAGCAGCGGGAAGGACAGGATGGGATATAAACTCAGGCA 780
DB 2131 AATCATCTATTCCTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAAACCCAGTC 2190
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCACTGTATGGAGCTCTCTGTTT 840
DB 2191 TTGAGCGGGAACCGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTCTGTTT 2250
QY 841 CACTCTATTTCACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCATGTTTCTTACGG 900
DB 2251 CATGCTATTTCACTCTATTAAATCTTCAACTGCACTCTTCTGGTCCATGTTTCTTACGG 2310
QY 901 CTCAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCCGTCACACCCGCT 960
DB 2311 CTGAGCTGAGCTTTTCGCTCGCATCCACACTGCTGTTTGGCCCGTCACACCCGCT 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGAGAGTGTCCACTGTCTCTGATCCAGCGAGGT 1020
DB 2371 GCTGACTTCCATCCCTTTGGATCCAGAGAGTGTCCACTGTCTCTGATCCAGCGAGGT 2430
QY 1021 ACCAATGGCAGCTCCCGATCAGGCTAAAGCTTGGCTTCTTCTGATGGCTAAGTGCC 1080
DB 2431 ACCAATGGCAGCTCCCGATCAGGCTAAAGCTTGGCTTCTTCTGATGGCTAAGTGCC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 2491 TGGGTTTGTCTTAATAGAACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550
QY 1141 CCACGGCTTCTTAATAGAGCTATAACACTCACCAGTGGCCCAAGATTCCTTCTGCTA 1200
DB 2551 CCACGGCTTCTTAATAGAGCTATAACACTCACCAGTGGCCCAAGATTCCTTCTGCTA 2609
QY 1201 TCTGTGAGGCCAAGAACCCCGAGGTGAGGAAGTGGAGGTGGCCACCATTTGGGAAG 1257
DB 2610 TCCATAGGCCAAGAACCCCGAGGTGAGGAAGTGGAGGTGGCCACCATTTGGGAAG 2666

RESULT 39
AF520507S2
LOCUS
DEFINITION
Homo sapiens individual 37 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.
AF520508
AF520508.1 GI:33410984
ACCESSION
VERSION
KEYWORDS
2 of 2
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2694)
REFERENCE

AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 37 allele B"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
43..57
/note="splice acceptor site"
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17576.1"
/db_xref="GI:33410985"
/translation="MALPYHIFLFTVLLPSFTLTAPPRCRWTSSSPYQBELWRQRP
GNIDAPSRSUSKGTPTTAHMPRNCYHSATLCMHANTHWGKMINPSCPLGV
TVCTYFTQTMGGGVQDQAREKHEVLSQLTRVHGTSSPYKGLDLXKLHETLR
HTRVLSFNFTLTGLHEVSAQNFCWICLFLNRPYVSIIPVQWNNFSTEINTSV
LVGPLVNLFTHTSNLCVRFSTNTYTNQCIRWVTPPTQIVCLPSGIFPVGCTSA
YRCLNGSESCFSLFVPMWTIVTEODLYSVISKPRNRPVILPFIAGVGLGALG
TGIGITSTOFYKLSQELNGDMERVADSLVTLQDOLNSLAUVLONRALDLTAE
RGGLCLFGECCYVNSGIVTEKVEIRDRIORRAHELNRTPGWLLSOWMFWILP
FLGPLAAILLGLLPGPCIFNLLVNFVSRISIEAVLQMEPRMQSKTKIYRRPLDRPASP
RSDVNDIKGTPEEISAAQPLRPNSAGSS"
1705..1720
misc_feature
/note="splice acceptor site"
LTR
1959..2694
ORIGIN
Query Match 77.8%; Score 1033.6; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 9.2e-304;
Matches 1132; Conservative 0; Mismatches 112; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTCTGCTAACCGCCAAAGAGAGGGGAACTGTTATTTT 60
DB 1423 TCAAAATCGAAGAGCTTTAGACTTCTGCTAACCGCTGAAAGAGGGGAACTGTTATTTT 1482
QY 61 AGGGAAGAACTCTTGTAGTATGTTAATCAATCTGGAATCATTAATCTGAGAAGTTAAAGA 120
DB 1483 AGGGAAGAACTCTGTTATGTTAATCAATTCGGAATCGTCACTGAGAAGTTAAAGA 1542
QY 121 AATTGAGATCGAATATAATGTAGACGAGAGGACCTTCAAAACACTGCACCCCTGGGGCT 180
DB 1543 AATTCGAGATCGAATACAACGTAGACGAGAGGAGCTTCGAAACACTGGACCCCTGGGGCT 1602
QY 181 CCTAGCAATGAGATGCGCTTGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
DB 1603 CCTAGCAATGAGATGCGCTTGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 1662
QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTCTCTTCTTAAAGTTTGTCTCTTCAGAAAT 300
DB 1663 GCTACTCTCTTTGGACCTGATCTTAACTCTCTTGTAACTTGTCTCTTCAGAAAT 1722
QY 301 TGAAGCTGTAAGCTTACAAATAGTTCTTTCAAATGGAAACCCAGATGCAGTCCATGACTAA 360
DB 1723 CGAAGCTGTAAACTA-----CAAATGGAGCCCAAGATGCAGTCCAAAGACTAA 1770

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCAGATCTGATGTTAATGACATCAA 1830
QY 421 AGTCACCCCTCCCGAGGAAATCTCAATGTCACAAACCCCTACTACACTCCAAATCAGTAGG 480
DB 1831 AGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGAGCAGTTGTGTCAGCCAACTCTCCCAACAGTACTTGGGTTTCTCTGTTAGA 540
DB 1891 AAGCAGTTAGAGCAGTCTGTCGGCCAACTCTCCCAACAGCATTAGGTTTCTCTGTTAGA 1950
QY 541 GGGTGGACTGAGACAGGAGTACTGCTGATTTCTTAGCTGACTAAGAAATCCNNAAGCT 600
DB 1951 TGGGGGACTGAGACAGGAGTACTGCTGATTTCTTAGCTGACTAAGAAATCCCTAAGCT 2010
QY 601 ANCTGGGAAGTGACCGATCCATCTTTTAAACATGGGGCTTGCACCTAGCTACACCCG 660
DB 2011 AGCTGGGAAGTGACCACTCCACCTTTTAAACAGGGGCTTGCAACTTAGCTACACCTG 2070
QY 661 ACCAATCAGAGCTCTACTAAATGCTTAATCAGGCAAAACAGGAGTAAAGCATTAGCC 720
DB 2071 ACCAATCAGAGCTCTACTAAATGCTTAATCAGGCAAAACAGGAGTAAAGCATTAGCC 2130
QY 721 AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAAGGATTGGGATATAAATCAGGCA 780
DB 2131 AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAATGATCGGATATAAATCAGGCA 2190
QY 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 840
DB 2191 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 2250
QY 841 CACTCTATTACCTCTATTAAATCATGCACTGCACTCTTCTGTCGGTGTGTTTTATGG 900
DB 2251 CATGCTATTACCTCTATTAAATCTTCACTGCACTCTTCTGTCGCAATGTTTCTTAGCG 2310
QY 901 CTCAAGCTGAGCTTTGTTGCGCACTCAACCACTGCTGTTTCCACCGCTCAGACGCCCT 960
DB 2311 CTGAGCTGAGCTTTGCTGCGCACTCAACCACTGCTGTTTCCCGCCACCGCAGACCC 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGAGAGTGCTCACTGTGCTCTGATCCAGCGAGGT 1020
DB 2371 GCTGACTTCCATCCCTCTGGATCATGAGGGTGTCCGCTGTGCTCTGATCCAGCGAGG 2430
QY 1021 ACCAATGCGCACTCCCGATCAGGCTAAAGGCTTGGCAATTTCTGTCATGCTAAGTCC 1080
DB 2431 ACCAATGCGCGCTCCCAATCGGGCTTAAAGGCTTGGCAATTTCTGTCATGCTAAGTCC 2490
QY 1081 TGGGTTTGTCTTAATGAACCTGAACTGGTCACTGGGTTCCATGGTTCTCTTCCATGAC 1140
DB 2491 TGGGTTTCACTAATTTAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGAC 2550
QY 1141 CCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTTGTA 1200
DB 2551 CCACAGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTT-GAA 2609
QY 1201 TCTGTGAGGCCCAAGAACCCCGAGGTTCAGAAAGTGGGCTTGGCCACCAATTTGGGAAG 1257
DB 2610 TCCATAAGGCCCAAGAACCCCGAGGTTCAGAAACACAGAGGCTTGGCCACCAATTTGGGAG 2666
RESULT 40
AF520511S2
LOCUS Homo sapiens individual 45 allele B, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520512
VERSION AF520512.1 GI:33410992
KEYWORDS 2 of 2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2694)
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G.,
Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
JOURNAL
PUBMED
14757826
2 (bases 1 to 2694)
Mallet, F., Bouton, O. and Oriol, G.
Direct Submission
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 45 allele B"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
misc_feature
43..57
/notes="splice acceptor site"
CDS
284..1900
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA017578.1"
/db_xref="GI:33410993"
/translation="MALPYHFLFTVLLPSFTLTAPPCKRMTSSPYGEFLWRMRP
GNIDAPYSRLSGTFTFTAHMPNRYHSATLCMHANTHYWTGMINPSPGGILV
TICWTYPTQMGSDGGVDOAREKHVEVISOLTRVHGTSPPYKGLDLSKLHETLT
HPELVSLNTLTGLHEVSAONPTNCWICLPLNFRPVSI PVPKNNFSTEINTSV
LVGPLVSNLEIHTSHLTCVKFNSNTYTTNSQIRWTPPTQIVCLPSGIFVCGTSA
YRLNGSSMSLEIFLVPMTIYTQDLISYVISKPRNKRFPILFVIGAGVLGALG
TGIGGTTTQFYKLSQELNDEMVRADSLVTLQDLNSLAVALVIONRRALLTAE
RGCTCLFLAGECCYVQNSQIVTEKVEIRDIQRAEELRNTGPWGLLSOMMPWLP
FLGLAAIILLLLFGPICFNLVNFVSSRIEAVKLQWPKMQSKTKIYRPLDRPASP
RSDVNDIKGTPPEISNAQPLLRPNAGSS"
misc_feature
1705..1720
/notes="splice acceptor site"
LTR
1959..2694
ORIGIN
Query Match 77.88; Score 1033.6; DB 9; Length 2694;
Best Local Similarity 90.14; Pred. No. 9.2e-304;
Matches 1132; Conservative 0; Mismatches 112; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGGAACCTGTTTATTTT 60
DB 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGGGGGAACCTGTTTATTTT 1482
QY 61 AGGGGAAGAATGCTGTTAGTATGTTTAACTCAATCTGGAATCATTTACTAGAGAAAGTTAAAGA 120
DB 1483 AGGGGAAGAATGCTGTTATGTTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542
QY 121 AATTGAGATCGAATATAGTACGAGAGGACCTTCAAAACACTGCACCCCTGGGGCCT 180
DB 1543 AATTGAGATCGAATATAGTACGAGAGGAGCTTCGAAACACTGGACCCCTGGGGCCT 1602
QY 181 CCTCAGCAATGATGTCCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
DB 1603 CCTCAGCAATGATGTCCTGGATTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1662
QY 241 TTTACTCCTCTTTGGACCCCTGTATCTTCACTCTGTTGTAAGTTTGTCTCTCCAGAAAT 300
DB 1663 GCTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTTAACTTGTCTCTCCAGAAAT 1722

Search completed: February 21, 2005, 11:52:42
Job time : 5718.48 secs

QY 301 TGAAGCTGTAAAGCTACAATAATAGTTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAA 360
DB 1723 CGAAGCTGTAAAGCTA-----CAAATGGAGCCCAAGATGCAGTCCCAAGACTAA 1770
QY 361 AATCTACCGTGAACCCCTGGACCGGCTGCTAGACTATCTCTGTATGTTTAAATGACATTTGA 420
DB 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCGCTGCTAGCCACCATCTGTATGTTAAATGACATCAA 1830
QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAATTTCACTAGG 480
DB 1831 AGGCACCCCTCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTTCACTAGG 1890
QY 481 AAGCAGTTAGACAGTTGTGACCAACCTCCCAACAGACTCTGGGTTTTCTCTGTTGAGA 540
DB 1891 AAGCAGTTAGACAGTTGTGACCAACCTCCCAACAGACTCTGGGTTTTCTCTGTTGAGA 1950
QY 541 GGTGGACTGAGAGACAGACTAGCTGGATTTTCTTAGGCTGACTAGAAATCCCNAAAGCCT 600
DB 1951 TGGGGGACTGAGAGACAGACTAGCTGGATTTTCTTAGGCTGACTAGAAATCCCTAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCCG 660
DB 2011 AGCTGGGAAGGTGACCAATCCATCTTTAAACAGGGGCTTGCNACTTAGCTCACACCTG 2070
QY 661 ACCAATCAGAGACTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
DB 2071 ACCAATCAGAGACTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGAGCAGCGGAAGAGCAAGGATTTGGGATATAAATCTCAGGCA 780
DB 2131 AATCATCTATTGCTGAGAGCAGCAGGAGGACAATGATCGGGATATAAACCAGATC 2190
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTT 840
DB 2191 TTGAGCCGCGAAGCAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTT 2250
QY 841 CACTCTATTTCATCTCTATTAAATCATGCAACTGCATCTTCTGGTCGTTGTTTTATGG 900
DB 2251 CATGCTATTTCATCTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGG 2310
QY 901 CTCAGCTAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGTCCACAGACCCGCT 960
DB 2311 CTTGAGCTGAGCTTTTCGCTCGCCATCCACACTGCTGTTTGGCCACCGCAGACCCACC 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCGAGGT 1020
DB 2371 GCTGACTCCATCCCTCTGGATCATCAGGGTGTCCGCTGTCTCTGATCCAGCGAGGC 2430
QY 1021 ACCCAATGGCACTCCCGATCAGGCTAAAGCTTTGCCATTTCTCTGATGGCTAAAGTGC 1080
DB 2431 ACCCAATGGCGCTCCCAATCGGCTAAAGCTTTGCCATTTCTCTGATGGCTAAAGTGC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCTAGCTGGTTCCATGTTCTCTTCCATGAC 1140
DB 2491 TGGGTTTCTCATCTAAATTTAGCTGAACACTAGTCTAGTCTGGTTTCCATGTTCTCTTGTGAC 2550
QY 1141 CCACGCTTCTAATAGAGCTATAAACACTACCCGATGGCCCAAGATTTCCATTTCTCTGCTGA 1200
DB 2551 CCACGCTTCTAATAGAGCTATAAACACTACCCGATGGCCCAAGTTCATTTCTTCTTCTT 2609
QY 1201 TCTGTGAGGCCAAGAACCCCGAGGTTCAGAGAAAGTGGGCTTGGCCACCATTTGGGAAG 1257
DB 2610 TCCATAAGGCCAAGAACCCCGAGGTTCAGAGAAACAGGAGCTTGGCCACCATCTTGGGAG 2666

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:25:25 ; Search time 685.525 Seconds
(without alignments)
11476.370 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatgaagagtttag.....gcaagatccccccagtaaca 1329

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1:	Geneseqn19808:*
2:	Geneseqn19808:*
3:	Geneseqn20008:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004as:*
13:	Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1326	99.8	1329	2 AAV43219	AAV43219 Multiple
2	1326	99.8	1329	2 AAX29704	Aax29704 Clone 5M6
3	1326	99.8	1329	12 ADG14849	Adg14849 MSRV asso
4	1299	97.7	1329	9 ADB84403	Adb84403 MSRV-1 as
5	1035.2	77.9	56093	6 ABL61744	Adl61744 Colon ade
6	1029.4	77.5	9502	10 ADF59718	Adf59718 Human con
7	1023.2	77.0	10499	3 ABN97929	Abn97929 Human ret
8	801.6	60.3	7466	5 AAS68626	Aas68626 DNA encod
9	798.2	60.1	21646	11 ACN4454	Acn4454 Human gen
10	778.8	58.6	1393	3 AAS31002	Aas31002 Human dia
11	775.4	58.3	2030	3 AAG3826	Aag3826 Nucleotid
12	774.2	58.3	7974	11 ACN88786	Acn88786 Breast ca
13	766.8	57.7	2074	6 AAD41225	Aad41225 Human EMB
14	747.6	56.3	2046	8 ACC46747	Acc46747 Human dit
15	731.8	55.1	2946	2 AAX77526	Aax77526 Human sec
16	731.8	55.1	2946	3 AAZ59468	Aaz59468 Human sec
17	731.8	55.1	2946	10 ADC38776	Adc38776 Human con
18	730.8	55.0	2930	6 AAD24195	Aad24195 Human syn
19	730.2	54.9	2781	5 AAF55630	Aaf55630 Nucleotid
20	728.6	54.8	6394	5 AAS84210	Aas84210 DNA encod

21	722.6	54.4	7582	2 AAX25665	Aax25665 Complete
22	722.6	54.4	7582	3 AAX59215	Aax59215 Human end
23	721.2	54.3	1136	2 AAX25660	Aax25660 Human end
24	721.2	54.3	1136	3 AAX59210	Aax59210 3' pol ge
25	712.6	53.6	2782	2 AAX25661	Aax25661 Human end
26	712.6	53.6	2782	3 AAX59211	Aax59211 5' non co
27	712.6	53.6	2782	5 AAH20069	Aah20069 HERV-W en
28	704.8	53.0	1894	4 AAI14608	Aai14608 Probe #45
29	704.8	53.0	1894	4 ABA56337	Aba56337 Human foe
30	704.8	53.0	1894	4 AAI35980	Aai35980 Probe #46
31	704.8	53.0	1894	4 ABA45822	Aba45822 Human bre
32	704.8	53.0	1894	4 ABA25978	Aba25978 Probe #44
33	704.8	53.0	1894	4 AAX30018	Aax30018 Human bon
34	704.8	53.0	1894	4 AAK04516	Aak04516 Human bra
35	704.8	53.0	1894	4 ABS29670	Abs29670 Human liv
36	704.8	53.0	1894	5 AAI04422	Aai04422 Probe #44
37	704.8	53.0	1894	6 ABS04589	Abs04589 Human gen
38	680.6	51.2	2942	5 AAS77313	Aas77313 DNA encod
39	651	49.0	893	5 AAS65964	Aas65964 DNA encod
40	643.4	48.4	46340	3 ABN97978	Abn97978 Human ret
41	643.4	48.4	161334	11 ACN44334	Acn44334 Human gen
42	639.8	48.1	783	2 AAX25669	Aax25669 Human end
43	637.8	48.0	780	7 ADS31352	Ads31352 Human gen
44	610.4	45.9	3464	10 ADE09587	Ade09587 Novel DNA
45	607	45.7	2784	3 ABN97930	Abn97930 Human ret

ALIGNMENTS

RESULT 1
AAV43219
ID AAV43219 standard; cDNA; 1329 BP.

AC AAV43219;

29-DEC-1998 (first entry)

Multiple sclerosis associated retrovirus fragment 6.

Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene; env gene; rheumatoid arthritis-associated virus; ss.

Multiple sclerosis associated retrovirus.

Key	Location/Qualifiers
CDS	2..490
FT	/tag= a
FT	/product= "Encodes protein AAW71069"
FT	/transl_except= (pos:77-79, appears to code for a stop codon)
FT	/transl_except= (pos:125-127, appears to code for a stop codon)
FT	/transl_except= (pos:137-139, appears to code for a stop codon)

WO9823755-A1.

04-JUN-1998.

26-NOV-1997; 97WO-IB001482.

26-NOV-1996; 96US-00756429.

(INMR) BIO MERIEUX.

Perron H, Beseme F, Bedin F, Paranhos-Baccala G; Komurian-Pradel F, Jollivet-Reynaud C, Mandrand B;

WPI; 1998-322732/28.

P-PSDB; AAW71069.

New nucleic acid from retroviruses - useful for diagnosis, prevention and

PT	treatment of, e.g. multiple sclerosis.			Db	781	TTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTTT	840
XX							
PS	Disclosure; Page 187-188; 286pp; English.			Qy	841	CACCTATTTCACCTCTATTAAATCATGCAATGCACTCTTCTGGTCCGTTTATATGG	900
XX							
CC	The present sequence represents a multiple sclerosis (MS) associated			Db	841	CACCTATTTCACCTCTATTAAATCATGCAATGCACTCTTCTGGTCCGTTTATATGG	900
CC	retrovirus (MSRV) genomic fragment used in the method of the invention.			Qy	901	CTCAAGCTGAGCTTTTGGTCCCATCCACACCTGCTGTTTGGCCACCGTACACAGCCCGCT	960
CC	The invention provides complete or partial genomic sequences of the MSRV-			Db	901	CTCAAGCTGAGCTTTTGGTCCCATCCACACCTGCTGTTTGGCCACCGTACACAGCCCGCT	960
CC	1 pol gene, gag gene and env gene, and polypeptides encoded by these			Qy	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGTCTCTGATCCAGCGAGGT	1020
CC	genes. The invention also provides antibodies raised against the			Db	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGTCTCTGATCCAGCGAGGT	1020
CC	polypeptides. The genomic sequences, polypeptides and antibodies are also			Qy	1021	ACCATTTGCCACTCCCGATCAGCTAAAGCTTGGCATTTGTTCTCCATGGCTAAGTGCC	1080
CC	claimed useful for diagnosing infection by MS and rheumatoid arthritis-			Db	1021	ACCATTTGCCACTCCCGATCAGCTAAAGCTTGGCATTTGTTCTCCATGGCTAAGTGCC	1080
CC	associated viruses, and also for prevention and treatment of infection			Qy	1081	TGGTGTGCTTAATAGAACTGAACACTGGTCACTGGGTTCATGTTCTTCCATGAC	1140
XX	with these viruses			Db	1081	TGGTGTGCTTAATAGAACTGAACACTGGTCACTGGGTTCATGTTCTTCCATGAC	1140
SQ	Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;			Qy	1141	CCACGGCTTCTAATAGAGCTATAAACAACCTCCGATGGCCCAAGATTCCATTCTTGTA	1200
	Query Match 99.8%; Score 1326; DB 2; Length 1329;			Db	1141	CCACGGCTTCTAATAGAGCTATAAACAACCTCCGATGGCCCAAGATTCCATTCTTGTA	1200
	Best Local Similarity 100.0%; Pred. No. 0;			Qy	1201	TCGTGAGGCCAAGAACCCAGGTCAGAGAANGTGAAGCTTGGCCACCATTTGGGAAGTGG	1260
	Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			Db	1201	TCGTGAGGCCAAGAACCCAGGTCAGAGAANGTGAAGCTTGGCCACCATTTGGGAAGTGG	1260
Qy	1	TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAAGAGGGGAACCTGTTATTTTT	60	Qy	1261	CCCACTGCCATTTTGGTAGCGGCCACACCATCTTGGAGCTGTGGAGCAAGGATCCC	1320
Db	1	TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAAGAGGGGAACCTGTTATTTTT	60	Db	1261	CCCACTGCCATTTTGGTAGCGGCCACACCATCTTGGAGCTGTGGAGCAAGGATCCC	1320
Qy	61	AGGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAGTTAAGA	120	Qy	1321	CCAGTAACA 1329	
Db	61	AGGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAGTTAAGA	120	Db	1321	CCAGTAACA 1329	
Qy	121	ATTTTGGATCCAAATATATGATGAGCAGAGGACCTTCAAAACACTGCACCTCGGGCCT	180				
Db	121	ATTTTGGATCCAAATATATGATGAGCAGAGGACCTTCAAAACACTGCACCTCGGGCCT	180				
Qy	181	CCTCAGCCAAATGGATGCCCTGGACTCTCCCTTTCTAGGACCTCTAGCAGCTATAATTT	240				
Db	181	CCTCAGCCAAATGGATGCCCTGGACTCTCCCTTTCTAGGACCTCTAGCAGCTATAATTT	240				
Qy	241	TTTACTCTCTTTGGACCCCTGATCTTCAACTTCTGTTAGTTTGTCTCTCCAGAAAT	300				
Db	241	TTTACTCTCTTTGGACCCCTGATCTTCAACTTCTGTTAGTTTGTCTCTCCAGAAAT	300				
Qy	301	TGAAGCTGTAAAGCTACAATATAGTTCTTCAAAATGGAAACCCAGATGCACTCACTAA	360				
Db	301	TGAAGCTGTAAAGCTACAATATAGTTCTTCAAAATGGAAACCCAGATGCACTCACTAA	360				
Qy	361	AATCTACCTGGACCCCTGGACCGGCTGTAGATCTGCTGTGATGTTAATGACATTGA	420				
Db	361	AATCTACCTGGACCCCTGGACCGGCTGTAGATCTGCTGTGATGTTAATGACATTGA	420				
Qy	421	AGTCACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG	480				
Db	421	AGTCACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG	480				
Qy	481	AAGCAGTTAGACAGTTGTGAGCCAACTCCCAACAGTACTTGGTGTCTGTTGAGA	540				
Db	481	AAGCAGTTAGACAGTTGTGAGCCAACTCCCAACAGTACTTGGTGTCTGTTGAGA	540				
Qy	541	GGGTGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCNAGCCT	600				
Db	541	GGGTGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCNAGCCT	600				
Qy	601	ANCTGGGAGGTGCACCGCATCCATCTTAAACATGGGCTTGCAACTTAGCTCACACCCG	660				
Db	601	ANCTGGGAGGTGCACCGCATCCATCTTAAACATGGGCTTGCAACTTAGCTCACACCCG	660				
Qy	661	ACCAATCAGAGAGCTCACTAAAATGCTAATCAGGCAAAAACAGGAGTTAAGCAATAGCC	720				
Db	661	ACCAATCAGAGAGCTCACTAAAATGCTAATCAGGCAAAAACAGGAGTTAAGCAATAGCC	720				
Qy	721	AATCATCTATTGCTGAGCAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA	780				
Db	721	AATCATCTATTGCTGAGCAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA	780				
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	840				

RESULT 2	
AAAX29704	standard; DNA; 1329 BP.
ID	AAAX29704
XX	
AC	AAAX29704;
XX	
DT	17-OCT-2003 (revised)
DT	27-AUG-2003 (revised)
DT	08-JUN-1999 (first entry)
XX	
DE	Clone 5M6 from MSRV-1.
XX	
KW	Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW	rheumatoid polyarthritis; ss.
XX	
OS	Viruses.
XX	
PN	FR2765588-A1.
XX	
PD	08-JAN-1999.
XX	
PF	07-JUL-1997; 97FR-00008816.
XX	
PR	07-JUL-1997; 97FR-00008816.
XX	
PA	(INMR) BIO MERIEUX.
XX	
DR	WPI; 1999-098275/09.
XX	
PT	P-PSDB; AAW99554.
XX	
XX	Nucleic acid sequences of retrovirus called MSRV-1 - associated with
XX	multiple sclerosis or rheumatoid polyarthritis.

CC arthritis. The methods and compositions of the present invention are
 CC useful for diagnosing, preventing and/or treating multiple sclerosis
 CC and/or rheumatoid arthritis. The present sequence is used in the
 CC exemplification of the invention.
 XX
 SQ Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;
 Query Match 99.8%; Score 1326; DB 12; Length 1329;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACTGTATTATTTT 60
 DB 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACTGTATTATTTT 60
 QY 61 AGGGGAAGAATGCTTTAGTATGTTAAATCAATCTGGAATCAATTAATGAGAAGTTAAAGA 120
 DB 61 AGGGGAAGAATGCTTTAGTATGTTAAATCAATCTGGAATCAATTAATGAGAAGTTAAAGA 120
 QY 121 AATTGAGATCGAATATATATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 DB 121 AATTGAGATCGAATATATATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 181 CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
 DB 181 CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
 QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTAAAGTTTGTCTTCCAGAAAT 300
 DB 241 TTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTAAAGTTTGTCTTCCAGAAAT 300
 QY 301 TGAAGCTTAAGCTTACAAATAGTTCTTCAATGGAACCCAGAGTGCAGTCCATGACTAA 360
 DB 301 TGAAGCTTAAGCTTACAAATAGTTCTTCAATGGAACCCAGAGTGCAGTCCATGACTAA 360
 QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTGATTAATGACATTGA 420
 DB 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTGATTAATGACATTGA 420
 QY 421 AGTACCCCTCCGAGGAATCTCACTGACAGACCCCTACTACTCACTCAATTCAGTAGG 480
 DB 421 AGTACCCCTCCGAGGAATCTCACTGACAGACCCCTACTACTCACTCAATTCAGTAGG 480
 QY 481 AAGCAGTTTAGAGCAGTTGTGACCAACCTCCCAACAGTACTTGGGTTTCTGTGAGA 540
 DB 481 AAGCAGTTTAGAGCAGTTGTGACCAACCTCCCAACAGTACTTGGGTTTCTGTGAGA 540
 QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCCTTAGGCTGACTAAGAATCCCAAGCCT 600
 DB 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCCTTAGGCTGACTAAGAATCCCAAGCCT 600
 QY 601 ANCTGGGAAGGTGACCGCATCATCTTTAAACATGGGCTTGCACACTAGCTCACACCG 660
 DB 601 ANCTGGGAAGGTGACCGCATCATCTTTAAACATGGGCTTGCACACTAGCTCACACCG 660
 QY 661 ACCAATCAGAGCTCAGTAAATGCTAATCAGGCAAAACAGAGGTAAAGCAATAGCC 720
 DB 661 ACCAATCAGAGCTCAGTAAATGCTAATCAGGCAAAACAGAGGTAAAGCAATAGCC 720
 QY 721 AATCATCTATTGCTTAGAGACACAGCGGAAGGACAGGATTGGGATATAAATCAGGCA 780
 DB 721 AATCATCTATTGCTTAGAGACACAGCGGAAGGACAGGATTGGGATATAAATCAGGCA 780
 QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 840
 DB 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 840
 QY 841 CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTATATGG 900
 DB 841 CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTATATGG 900
 QY 901 CTCAGCTGAGCTTTTGTTCGCAATCCACCACTGCTGTTTGGCCACCGTFCAGACCCGCT 960
 DB 901 CTCAGCTGAGCTTTTGTTCGCAATCCACCACTGCTGTTTGGCCACCGTFCAGACCCGCT 960

Db 901 CTCAGCTGAGCTTTTGTTCGCAATCCACCACTGCTGTTTGGCCACCGTFCAGACCCGCT 960
 QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGGAGGT 1020
 DB 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGGAGGT 1020
 QY 1021 ACCCATTCGCACTCCGATCAGGCTAAAGGCTTCCCATTTGTTCTGCTGATGCTTAAGTCCC 1080
 DB 1021 ACCCATTCGCACTCCGATCAGGCTAAAGGCTTCCCATTTGTTCTGCTGATGCTTAAGTCCC 1080
 QY 1081 TGGGTTTTGCTCTAATAGAACTGAACACTGCTGGTTCATGGTTCATTTCTTCCATGAC 1140
 DB 1081 TGGGTTTTGCTCTAATAGAACTGAACACTGCTGGTTCATGGTTCATTTCTTCCATGAC 1140
 QY 1141 CCAGGCTTCTAATAGAGCTATAACACTCACCAGATGGCCCAAGATTCCATTTCTTGSTA 1200
 DB 1141 CCAGGCTTCTAATAGAGCTATAACACTCACCAGATGGCCCAAGATTCCATTTCTTGSTA 1200
 QY 1201 TCTGTGAGGCCAAGAACCCAGAGTCCAGAAAGTGGCTTGGCCACCATTTTGGGAAGTGG 1260
 DB 1201 TCTGTGAGGCCAAGAACCCAGAGTCCAGAAAGTGGCTTGGCCACCATTTTGGGAAGTGG 1260
 QY 1261 CCCACTGCCATTTTGGTAGCGGCCACCACTCTTGGAGCTGTGGAGCAAGATCCC 1320
 DB 1261 CCCACTGCCATTTTGGTAGCGGCCACCACTCTTGGAGCTGTGGAGCAAGATCCC 1320
 QY 1321 CCAGTAACA 1329
 DB 1321 CCAGTAACA 1329
 RESULT 4
 ADB84403
 ID ADB84403 standard; DNA; 1329 BP.
 XX
 AC ADB84403;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE MSRV-1 associated DNA sequence #21.
 XX
 KW ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
 KW reverse transcriptase; ribonuclease H.
 XX
 OS Unidentified;
 XX
 FN US2003039664-A1.
 XX
 PD 27-FEB-2003.
 XX
 PF 26-NOV-1997; 97US-00979847.
 XX
 PR 26-NOV-1996; 96US-00756429.
 XX
 PA (PERR/) PERRON H.
 PA (BESE/) BESEME F.
 PA (BEDI/) BEDIN F.
 PA (PARA/) PARANHOS-BACCALA G.
 PA (KOMU/) KOMURIAN-PRADEL F.
 PA (JOLI/) JOLIVET-REYNAUD C.
 PA (MAND/) MANDRAND B.
 PA (GARS/) GARSON J A.
 PA (TUKU/) TUKU P W.
 XX
 PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
 PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
 XX
 DR WPI; 2003-512253/48.
 XX
 PT New isolated or purified nucleic acid associated with multiple sclerosis
 PT and/or rheumatoid arthritis, useful for detecting a virus associated with
 PT multiple sclerosis or rheumatoid arthritis in a biological sample.
 XX

Claim 31; Page 80; 193pp; English.

PS The invention relates to an isolated or purified nucleic acid from a
 XX virus associated with multiple sclerosis and/or rheumatoid arthritis,
 CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
 CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
 CC proteins or defined peptides (including immunodominant peptides,
 CC antigenic peptides or conserved motifs). Also included are a process for
 CC detecting a virus associated with multiple sclerosis or rheumatoid
 CC arthritis in a biological sample, a nucleic acid probe for the detection
 CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
 CC primer for the amplification by polymerization of a nucleic acid of a
 CC viral material associated with multiple sclerosis or rheumatoid
 CC arthritis, a polypeptide exhibiting an inhibitory activity on the
 CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
 CC and an antibody directed against the MSRV-1 virus obtained by
 CC immunologically reacting a human or animal body or cells with an
 CC immunogenic agent consisting of the antigenic polypeptide defined above.
 CC The nucleic acids are useful for detecting a biological sample, a virus
 CC associated with multiple sclerosis or rheumatoid arthritis, or for
 CC detecting in a biological sample, the presence of or exposure to a virus
 CC associated with multiple sclerosis or rheumatoid arthritis. The present
 CC sequence is a claimed MSRV-associated sequence whose identity cannot be
 CC accurately determined. Note: The SEQ ID numbers for the sequences as
 CC displayed in the main body of the patent do not match the SEQ ID numbers
 CC in the sequence listing. Consequently those sequences mentioned in the
 CC claims may not be the sequences the authors intended to claim.

XX SQ Sequence 1329 BP; 339 A; 346 C; 282 G; 332 T; 0 U; 30 Other;

Query Match		97.7%	Score 1299;	DB 9;	Length 1329;
Best Local Similarity		98.0%;	Pred. No. 0;		
Matches 1302;		Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;
QY	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAAACCTGTTTATTTT	60		
DB	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAAACCTGTTTATTTT	60		
QY	61	AGGGAGAGATCGTGTAGTATGTTAATCAATCTGGAATCACTTACCTGAGAAAGTTAAAGA	120		
DB	61	AGGGAGAGATCGTGTAGTATGTTAATCAATCTGGAATCACTTACCTGAGAAAGTTAAAGN	120		
QY	121	AATTTGAGATCGAATATAATGTAGACGAGGAGACCTTCAAAACACCTGCGCCCTGGGGCCT	180		
DB	121	AATTTGAGATCGAATATAATGTAGACGAGGAGACCTTCAAAACACCTGCGCCCTGGGGCCN	180		
QY	181	CCTCAGCCAAATGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATAT	240		
DB	181	CCTCAGCCAAATGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATN	240		
QY	241	TTTACTCTCTTTGGACCCGTATCTCACTTCTTGTAGTTTGTCTCTCCAGAAAT	300		
DB	241	TTTACTCTCTTTGGACCCGTATCTCACTTCTTGTAGTTTGTCTCTCTCCAGAA	300		
QY	301	TGAAGCTGTAAAGCTCAAAATAGTCTTCAATGGAACCCAGATGCGATGCCATGACTAA	360		
DB	301	TGAAGCTGTAAAGCTCAAAATAGTCTTCAATGGAACCCAGATGCGATGCCATGACTAN	360		
QY	361	AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGN	420		
DB	361	AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGN	420		
QY	421	AGTCACCCCTCCCGAGGAATCTCAACTGACACACCCCTACTACCTCCAATTCAGTAGG	480		
DB	421	AGTCACCCCTCCCGAGGAATCTCAACTGACACACCCCTACTACCTCCAATTCAGTAGN	480		
QY	481	AAGCAGTTAGCAGTTGTTCAGGCAACCTCCCAACAGACTTTGGGTTTCTCTGTTGAGA	540		
DB	481	AAGCAGTTAGCAGTTGTTCAGGCAACCTCCCAACAGACTTTGGGTTTCTCTGTTGAGN	540		
QY	541	GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAGCCCT	600		
DB	541	GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAGCCN	600		

QY	601	ANCTGGAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660
DB	601	ANCTGGAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCN	660
QY	661	ACCAATCAGAGAGCTCATAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	720
DB	661	ACCAATCAGAGAGCTCATAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCN	720
QY	721	AATCATCTATTGCTGAGAGCACAGGGAAGCAAGGATTGGGATATAAACTCAGGCA	780
DB	721	AATCATCTATTGCTGAGAGCACAGGGAAGCAAGGATTGGGATATAAACTCAGGCN	780
QY	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTT	840
DB	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTN	840
QY	841	CACTCTATTTCATCTCTATTAAATCATGCACTGCACTCTCTTGGTCCGTTTATGN	900
DB	841	CACTCTATTTCATCTCTATTAAATCATGCACTGCACTCTCTTGGTCCGTTTATGN	900
QY	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCCGTCACAGCCGCT	960
DB	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCCGTCACAGCCGCGN	960
QY	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAGGT	1020
DB	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAGNN	1020
QY	1021	ACCCATTTGCCATCCCGATCAGGCTTAAGGCTTGCCATTTGCTGATGGCTTAAGTGC	1080
DB	1021	ACCCATTTGCCATCCCGATCAGGCTTAAGGCTTGCCATTTGCTGATGGCTTAAGTGN	1080
QY	1081	TGGGTTTGTCTTAATAGAACTGAACTGCTCACTGGGTTCATGGTCTCTTCCATGAC	1140
DB	1081	TGGGTTTGTCTTAATAGAACTGAACTGCTCACTGGGTTCATGGTCTCTTCCATGNN	1140
QY	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATCTCTTGGTA	1200
DB	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATCTCTTGGNN	1200
QY	1201	TCTGTGAGCCCAAGAACCCAGGTCAGAGAANGTGGGCTTGCACCATTTGGGAAGTGG	1260
DB	1201	TCTGTGAGCCCAAGAACCCAGGTCAGAGAANGTGGGCTTGCACCATTTGGGAAGTNN	1260
QY	1261	CCCACTGCCATTTTGTAGCGGCCACCACTTGGGAGCTGTGGGAGCAAGGATCCC	1320
DB	1261	CCCACTGCCATTTTGTAGCGGCCACCACTTGGGAGCTGTGGGAGCAAGGATCNN	1320
QY	1321	CCAGTAACA 1329	
DB	1321	CCAGTAACA 1329	

RESULT 5
 ABL61744
 ID ABL61744 standard; DNA; 56093 BP.
 XX
 AC ABL61744;
 XX
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Colon adenocarcinoma related gene sequence SEQ ID NO:81.
 XX
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 XX gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX

PD 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-02331133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 22-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 28-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 81; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,

CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;
SQ Query Match 77.9%; Score 1035.2; DB 6; Length 56093;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGCTAACCGCAAAAGAGGGGGAACCTGTTTATTTT 60
Db 37018 TCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTT 37077
Qy 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGATCATTTACTCAGAAAGTTAAAGA 120
Db 37078 AGGGGAAGATGCTGTTAGTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 37137
Qy 121 AATTTGAGATCGAATATAATGTAGAGCAGAGGAGCTTCAAAAACACTGCACCTGGGCGCT 180
Db 37138 AATTCGAGATCGAATACAACTGAGAGCAGAGGAGCTTCGAAACACTGCAGCTGGGCGCT 37197
Qy 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
Db 37198 CCTCAGCAATGGATGCCCTGGATCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 37257
Qy 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTTCTTAAAGTTTCTCTTCCAGAAAT 300
Db 37258 GCTACTCTCTTTGGACCCCTGTATCTTAACTCTTCTTAACTTTTCTTCTTCCAGNAT 37317
Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db 37318 CGAAGCTGTAAGCTA-----CAAAATGGAGCCCAAGATGCAGTCCAAAGACTAA 37365
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db 37366 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCCGATCTGATGTTAATGACATCAA 37425
Qy 421 AGTCACCCCTCCGAGGAAATCTCACTGCACACCCCTACTACACTCCAAATTCAGTAGG 480
Db 37426 AGGCACCCCTCTCGAGGAATCTCAGCTGCACACCTCTACTAGCCCAATTCAGCAGG 37485
Qy 481 AAGCAGTTAGACGAGTTGTGACGCCAACCTCCCCAACAGTACTTGGGTTTTCTGTTAGA 540
Db 37486 AAGCAGTTAGACGCTGCTGGCCAACTCTCCCAACAGCACTAGGTTTTCTGTTAGA 37545
Qy 541 GGGTGGACTGAGACAGGACTAGTGGATTTCTTAGCTGACTAAGAAATCCNAAGCCT 600
Db 37546 TGGGGGACTGAGACAGGACTAGTGGATTTCTTAGCTGACTAAGAAATCCCTAAGCCT 37605
Qy 601 ANCTGGGAAGGTGACCGCATCTTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660
Db 37606 AGCTGGGAGGTGACCATCTTAAACATGGGGCTTGCAACTTAGCTCACACCTG 37665
Qy 661 ACCAATCAGAGCTCTCTAAATGCTAATCAGGCAAAACAGGAGGTTAAGCAATAGCC 720
Db 37666 ACCAATCAGAGCTCTCTAAATGCTAATCAGGCAAAACAGGAGGTTAAGCAATAGCC 37725
Qy 721 AATCATCTATTGCTGAGACAGCGGAGGACAGGATTTGGGATATAAATCTCAGCA 780
Db 37726 AATCATCTATTGCTGAGACAGCGGAGGACAGGATTTGGGATATAAATCTCAGCA 37785
Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCATTTGATGGAGCTCTGTTTT 840
Db 37786 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCATTTGATGGAGCTCTGTTTT 37845
Qy 841 CACTCTATTTCATCTCTAATAATCATGCAACTGCACTCTTCTGGTCCCTGTTTTATGG 900
Db 37846 CATGCTATTTTCACTCTATTAAATCTTGAACACTGCACTCTTCTGGTCCATGTTTCTACGG 37905
Qy 901 CTCAGCTGAGCTTTGTTGCGCACTCCACCACTGCTGTTTGCACCGTCACAGACCCGCT 960

37906 CTTGAGCTGAGCTTTTCGCTCGCATCCACCACTGCTGTTTTCGCGCCACGCGAGCCGCGC 37965
 961 GTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGCTGATCCAGCGAGT 1020
 37966 GTGACTCCCATCCCTCTGGATCATGACAGGTTGTCGCTGTGCTCTGATCCAGCGAGC 38025
 1021 ACCATTGGCACTCCCGATCAGCTAAAGCTTGGCCATTGTTCCCTGATGGCTAAAGTGC 1080
 38026 ACCATTGGCGCTCCCAATCGGGCTAAAGCTTGGCCATTGTTCCCTGATGGCTAAAGTGC 38085
 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTTCCATGAC 1140
 38086 TGGGTTTCATCTTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGAC 38145
 1141 CACGGCTTCTAATAGACTATTAACACTACCGCATGGCCCAAGATTCCATTCTCTTGGTA 1200
 38146 CCACAGCTTCTAATAGACTATAACACTACCGCATGGCCCAAGATTCCATTCTCTT-GAA 38204
 1201 TCTGTGAGCCCAAGAACCCAGGTCAAGAAAGTGGGCTTGGCCACATTTGGGAAG 1257
 38205 TCATTAAGCCCAAGAACCCAGGTCAAGAAAGTGGGCTTGGCCACATTTGGGAG 38261

RESULT 6

ID ADF59718 standard; cDNA; 9502 BP.

XX AC

XX AC

XX AC

DT 12-FEB-2004 (first entry)

XX AC

XX AC

DE Human contig polynucleotide sequence SEQ ID NO:2085.

KW biological activity; genetic engineering; hybridisation probe; oligomer;
 KW primer; chromosome mapping; gene mapping; recombinant protein production;
 KW human; gene; ss.

XX Homo sapiens.

XX WO2003080795-A2.

XX PD 02-OCT-2003.

XX PF 09-AUG-2002; 2002WO-US025485.

XX PR 09-AUG-2001; 2001US-0311261P.

XX PA (HYSE-) HYSSEQ INC.

XX PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX WPI; 2003-876918/81.

XX DR P-PSDB; ADF60170.

XX New polynucleotides, useful as hybridization probes, oligomers or
 PT primers, for chromosome or gene mapping, for the recombinant production
 PT of proteins, and for generating antisense DNA or RNA.

XX Example 2; SEQ ID NO 2085; 571bp; English.

XX The present invention describes isolated polynucleotide sequences (I),
 CC which encode polypeptides (II) with biological activity. Also described:
 CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)
 CC a host cell genetically engineered to comprise (I) which is operatively
 CC associated with a regulatory sequence that modulates expression of (I) in
 CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition
 CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed
 CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of
 CC (4) in a sample; (8) identifying a compound that binds to the polypeptide
 CC of (4); (9) producing the polypeptide of (4); and (10) a collection of
 CC polynucleotides comprising at least one of the polynucleotide sequences
 CC (I). The polynucleotides (I) can be used as hybridisation probes,
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant

CC production of proteins, and for generating antisense DNA or RNA. The
 CC present sequence represents a human contig polynucleotide sequence, which
 CC is used in an example from the present invention.

XX Sequence 9502 BP; 2813 A; 2433 C; 1992 G; 2263 T; 0 U; 1 Other;
 SQ

Query Match 77.5%; Score 1029.4; DB 10; Length 9502;
 Best Local Similarity 87.9%; Pred. No. 0;
 Matches 1178; Conservative 0; Mismatches 129; Indels 33; Gaps 4;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAAGAGGGGAACTGTTATTTT 60
 Db 2575 TCAAAATCGAAGAGCTTTAGACTTCTTAACCGCTGAAAGAGGGGAACTGTTATTTT 2634
 QY 61 AGGGAAGAATGCTCTAGTATGTTAAATCAATCTGGAATCATTTACTAGAAAAGTTAAAGA 120
 Db 2635 AGGGAAGAATGCTCTAGTATGTTAAATCAATCCGGAATCTGCTAGAAAAGTTAAAGA 2694
 QY 121 AATTTGAGATCGAATATATTAATGTAGAGCAGAGGACCTTCAAAAACACTGCAACCTTGGGGCCT 180
 Db 2695 AATTCGAGATCGAATATCAACGTTAGAGCAGAGGAGCTTCGAAACACTGGACCTTGGGGCCT 2754
 QY 181 CCTCAGCAATGAGTGCCTTGGACTCTCCCTTTCTTAGGACCTCTAGAGCTATATATT 240
 Db 2755 CCTCAGCAATGAGTGCCTTGGACTCTCCCTTTCTTAGGACCTCTAGAGCTATATATT 2814
 QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTAAAGTTTGTCTTCTTCCAGAAT 300
 Db 2815 GCTACTCTCTTTGGACCTGATCTTCAACTTCTTGTAAAGTTTGTCTTCTTCCAGAAT 2874
 QY 301 TGAAGCTGTAAAGCTTCAAAATAGTTTCTTCAAAATGGAACCCAGATGTCAGTCCATGACTAA 360
 Db 2875 CGAAGCTGTAAAGCTTCAAAATGGAACCCAGATGTCAGTCCATGACTAA 2922
 QY 361 AATCTACCGTGAACCCCTGGACCGGCTGTAGACTATGCTCTGTATGTTAAAGTGAATGA 420
 Db 2923 GATCTACCGCAGACCCCTGGACCGGCTGTAGACTATGCTGTATGTTAAAGTGAATGA 2982
 QY 421 AGTCAACCTTCCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAAATTCAGTAGG 480
 Db 2983 AGGCACCCCTCTCTGAGGAATCTCAACTGCAACCCCTACTACACTCCAAATTCAGTAGG 3042
 QY 481 AAGCAGTTAGAGCAGTTGTGACGCAACCTTCCCAACAGTACTTGGGTTTTCTTGTGAGA 540
 Db 3043 AAGCAGTTAGAGCAGTTGTGACGCAACCTTCCCAACAGTACTTGGGTTTTCTTGTGAGA 3102
 QY 541 GGGTGAAGTGAAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCT 600
 Db 3103 TGGGGGAGCTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCT 3162
 QY 601 ANCTGGGAGGTGACCGCATCCATCTTTAAACATGGGCTTGCACCTTAGCTCACAACCCG 660
 Db 3163 AGCTGGGAAGGTAAACCATCATCCCTTTAAACATGGGCTTGCACCTTAGCTCACAACCCG 3222
 QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAA 711
 Db 3223 ACCAATCAGTAGTAAGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAA 3282
 QY 712 GCAATAGCCAAATCATCTATTGCTGAGACACAGCGGGAAGGACAGGATTTGGATATAA 771
 Db 3283 GAAATAGCCAAATCATCTATTGCTGAGACACAGCGGGAAGGACAGGATTTGGATATAA 3342
 QY 772 ACTCAGGCAATTCAGCCAGCAACGACCCCTTTGGGCTCCCTCCCTTGTATGGGAG 831
 Db 3343 ACCCAGGCAATTCAGCCAGCAACGACCCCTTTGGGCTCCCTCCCTTGTATGGGAG 3402
 QY 832 CTCTGTTTTCACTCTATTTCACCTATTAAATCATGCAACTGCA--CTCTTCTGGTCCGT 889
 Db 3403 CTCTGT-----TTCCACTCTATTAAATCTTGCACCTGCACCTCTCTTCTGGTCCGT 3452
 QY 890 GTTTTTATGGTCAAGCTGAGCTTTTGTTCGCCATTCACCACTGCTGTTTGGCCACCGTC 949
 Db 3453 GTTTGTTACGGATTGAGCTGAGCTTTTGTTCGCCATTCACCACTGCTGTTTGGCCGCTGT 3512

QY 950 ACAGACCGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGA 1009
 DB 3513 GCAGACCTGCCACTGACTTCCATCCCTCTGGATCCAGGGGTGTCGCTGTGCTCTGA 3572
 QY 1010 TCCAGCGAGGTACCCATTGCCACTCCCGATCAGGCTAAAGGCTTGCCATTGTTCTCTGCAT 1069
 DB 3573 TCCAGCGAGGCGCCATTGCCCTCCCGATCGGGCTAAAGGCTTGCCATTGTTCTCTGCAC 3632
 QY 1070 GGCTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCTACTGGTTCCATGGTTC 1129
 DB 3633 GGCTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCTACTGGTTCATGGTTC 3692
 QY 1130 TCTTCATGACCCAGCGCTTCTTAATAGAGCTATTAACACTCACCGCATGCCGCCAAGATTCC 1189
 DB 3693 TCTTCGTGACCCAGCGCTTCTTAATAGAACTATTAACACTTACCACTATGCCCAAGATTCC 3752
 QY 1190 ATTCTCTGTATCTGTAGGCGCAAGACCCAGGTTCAGAGAAAGTGGAGGCTTGCCACCAT 1249
 DB 3753 ATTCTCTGTAAATCCGTGAGGCGCAAGAACTCCAGGTCAGAGAAATACGAGGCTTGCCACCAT 3812
 QY 1250 TTGGGAAGTGGCCCACTGCCATTTTGGTATGCGGCCCCACCACTCTTTGGGAGCTGTGGGA 1309
 DB 3813 CTGGGAAGCGGCTGCTACTCGCTTTGGAAAGTGGTTCCACCACATCTTGGGAGCTCTGTGA 3872
 QY 1310 GCAAGGATCCCCAGTAACA 1329
 DB 3873 GCAAGGACCCCGGTAAACA 3892

RESULT 7

ABN97929 ID ABN97929 standard; DNA; 10499 BP.
 XX AC ABN97929;
 XX DT 01-AUG-2002 (first entry)
 XX DE Human retroviral sequence HERV-7q.
 XX KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
 XX KW multiple sclerosis; ds.
 XX OS Human endogenous retrovirus.
 XX PN WO967395-Al.
 XX PD 29-DEC-1999.
 XX PF 23-JUN-1999; 99WO-FR001513.
 XX PR 23-JUN-1998; 98FR-00007920.
 XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX PI Alliel PM, Perin J, Rieger F;
 XX WPI; 2000-160587/14.
 XX PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
 XX PT for diagnosis, treatment and prevention of autoimmune and neurological
 XX PT diseases.
 XX PS Claim 3; Fig 1; 225pp; French.
 XX CC The present invention relates to new nucleic acid sequences of human
 XX CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
 XX CC Regulatory elements associated with HERV-7q may alter expression of other
 XX CC genes (even remote genes) on the same chromosome, inducing immunological
 XX CC and/or neurological changes (which may be pathological or protective/
 XX CC curative). HERV-7q peptides can be used to improve efficiency of the
 XX CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
 XX CC sequences can be used in immunogenic or vaccinating compositions, for

CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention
 XX SQ Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 U; 0 Other;
 Query Match 77.0%; Score 1023.2; DB 3; Length 10499;
 Best Local Similarity 90.1%; Pred. No. 0;
 Matches 1132; Conservative 0; Mismatches 111; Indels 14; Gaps 3;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTGTTATTTT 60
 DB 9018 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGAACTGTGTTATTTT 9077
 QY 61 AGGGGAAGAAATGCTGTGTATGTTTAAATCAATCTGGAATCATTTACTGAGAAAGTTAAGA 120
 DB 9078 AGGGGAAGAAATGCTGTGTATGTTTAAATCAATCTGGAATCGTCACTGAGAAAGTTAAGA 9137
 QY 121 AATTTGAGATCGAATAATATGTAGACAGAGGACCTTCAAAAACACTGCACCTGGGGCCT 180
 DB 9138 AATTCGAGATCGAATAACAGTATAGACAGAGGAGCTTCGAAACACTGGAACCTGGGGCCT 9197
 QY 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTTTATAGGACCTCTAGCAGCTATAATATT 240
 DB 9198 CCTCAGCAATGGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 9257
 QY 241 TTTACTCTCTTTGGACCCCTGTATCTCAACTCTCTTAAAGTTTCTCTCTCCAGAAAT 300
 DB 9258 GCTACTCTCTTTGGACCCCTGTATCTTAACTCTCTTAACTTTTCTCTCTCCAGAAAT 9317
 QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAGTCCATGACTAA 360
 DB 9318 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCAGTCCAGACTAA 9365
 QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTGTAGACTATGCTCTGATTTAATGACATTGA 420
 DB 9366 GATCTACCGGAGACCCCTGGACCGGCTGTGTAGCCACGATCTGATGTTAATGACATCAA 9425
 QY 421 AGTCACCCCTCCGAGGAAATCTCAATGTCACACACCCCTACTACACTCCAATCCAGTAGG 480
 DB 9426 AGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACTCTACTAGCCCAATTCAGCAGG 9485
 QY 481 AAGCAGTTAGACAGTGTGTGAGCCAACTCTCCCAACAGTACTTGGGTTTTCTCTGTGAGA 540
 DB 9486 AAGCAGTTAGAGC--GGTCTCGGCCAACCTCCCAACAGCACTTAGTGGTTTTCTGTGAGA 9544
 QY 541 GGGTGGACTGAGACAGAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCAAGCCT 600
 DB 9545 TGGGGGACTGAGACAGAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCTAAGCCT 9604
 QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660
 DB 9605 AGCTGGGAAGGTGACCGCATCCATCTTTAAACAGCGGGCTTGCAACTTAGCTCACACCTG 9664
 QY 661 ACCAATCAGAGGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC 720
 DB 9665 ACCAATCAGAGGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC 9724
 QY 721 AATCATCTATTGCTGAGAGCAGCAGCGGAAAGGACAGGATGGGATATAAATCAGGCA 780
 DB 9725 AATCATCTATTGCTGAGAGCAGCAGCGGAGGCAATGATCGGGATATAAATCAGGCA 9784
 QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
 DB 9785 TTCAAGCCAGCAACAGGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 9844
 QY 841 CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTATGG 900
 DB 9845 CATGCTATTTCATCTATTAAATCTTCAACTGCACTCTTCTGGTCCATGTTTCTTACGG 9904
 QY 901 CTCAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTTGCCACCGTCAGACAGCCGCT 960


```
QY 1013 AGCAGGTACCATTCCTCCACTCCCGATCAGCTAAAGCTTGCCATTGTTCTCTGATGGC 1012
DB |||||
709 AGAGAGGCACCAATTCCTGATTCCTGATGGGCTAAAGGCTTGCCATTGTTCTCTGANGAC 768
QY 1073 TAAGTGCTGGTGTGCTTAATAGAACTGAACACTGTGTCACTGGGTTCCCATGTTCTCT 1132
DB |||||
769 TAAGTGCCGGGTTCACTTAATCGAGCTGAACACTAGTCTGCTGGTTCACGATCTCT 828
QY 1133 TCCATGACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCAT 1192
DB |||||
829 TCCGTGACCCACGACTTCTAATAGAGCTATAACACTCACCGCACGGCCCAAGATTCAT 888
QY 1193 CTTGCTATCTGTAGGCGCAAGAACCCAGCTCAGAGAACTGAGGCTTGCCACCATTTG 1252
DB |||||
889 CTTTCGAATCCGTGAGGCCAAGAACCCAGGTCAGAGAACTGAGGCTTGCCACCATTT 948
QY 1253 GGAAGTGCCCACTGCCATTTTGGTAGCGGCCACCACTCTGGGAGCTGTGGAGCA 1312
DB |||||
949 GGAAGTGCCCTGCGGCCATTTTGGNAGTGGCTGCCACCATCTTGGGAGCTCTGGGAGCA 1008
QY 1313 AGGATCCCCCAGTAACA 1329
DB |||||
1009 AGGACCCCTGTGTACA 1025

RESULT 11
AAA63826
ID AAA63826 standard; DNA; 2030 BP.
XX
AC AAA63826;
XX
DT 06-AUG-2003 (revised)
DT 04-DEC-2000 (first entry)
XX
XX
DE Nucleotide sequence of the MSRV-1 3' env and LTR regions.
XX
KW MSRV-1; pol region; long terminal repeat; LTR; RUS region; retrovirus;
KW ss.
XX
OS Multiple sclerosis associated retrovirus.
XX
PH Key Location/Qualifiers
FT CDS 1..1629
FT /tag= a
FT /note= "Contains one termination codon"
FT
FT sig_peptide 1..81
FT /tag= b
FT CAAT_signal 1800..1807
FT /tag= c
FT CAAT_signal 1858..1864
FT /tag= d
FT TATA_signal 1906..1911
FT /tag= e
FT polyA_signal 1996..2002
FT /tag= f
XX
FN WO200047745-A1.
XX
XX 17-AUG-2000.
XX
XX 15-FEB-2000; 2000WO-IB0000159.
XX
XX 15-FEB-1999; 99EP-00420041.
XX (INNR ) BIO MERIEUX.
XX
XX Paranhos-Baccala G, Perron H, Komurian-Pradel F;
XX
XX WPI; 2000-506097/45.
XX
XX P-PSDB; AAB08195.
XX
XX Nucleotide fragment of LTR-RUS region from Multiple Sclerosis retrovirus
```

```
PT (MSRV) used to detect the presence of MSRV-1 retrovirus in a biological
PT sample.
XX
XX Disclosure; Fig 2; 23pp; English.
XX
CC The present sequence represents the nucleotide sequence corresponding to
CC the 3' env region and long terminal repeat sequences from clone CL6 of
CC Multiple Sclerosis retrovirus (MSRV-1). The specification describes a
CC long terminal repeat (LTR)-RUS region which encodes the expression of a
CC MSRV-1 protein. This is unusual for LTRs, in particular in the RUS
CC region. The sequence includes CAAT and TATA signals which are present in
CC the U3 and R regions and are not directed towards the CPS indicated in
CC the features table. Probes and antibodies to the MSRV-1 retrovirus
CC protein and encoding polynucleotide sequences are used to detect the
CC presence of MSRV-1 retrovirus in a biological sample. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 2030 BP; 574 A; 559 C; 387 G; 510 T; 0 U; 0 Other;
Query Match 58.3%; Score 775.4; DB 3; Length 2030;
Best Local Similarity 92.8%; Pred. No. 2.1e-243;
Matches 812; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 1 TCAAAATCGAAGAGCTTTAGACTTTGTAACCGCCAAAGAGGGGAACTCTTTATTTT 60
DB |||||
1140 TCAAAATCGAAGAGCTTTAGACTTTGTAACCGCCAAAGAGGGGAACTCTTTATTTT 1199
QY 61 AGGGGAAGATGCTGTAGTATGTTAATCAATCTGGATCACTTATGAGAAAGTAAAGA 120
DB |||||
1200 AGGAGAAAGACGCTGTTATTATGTTAATCAATCAGAAATGTCATCGAAGAAAGTAAAGA 1259
QY 121 AATTGTGATCGAATAATAATGTAGAGCAGAGAGCTTCAAAACACTGCACCTCGGGCCT 180
DB |||||
1260 AATTGTGATCGAATAATAATGTAGAGCAGAGAGCTTCAAAACACTGCACCTCGGGCCT 1319
QY 181 CCTCAGCAATGGATGCCCTGGAGCTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATT 240
DB |||||
1320 CCTCAGCAATGGATGCCCTGGAGCTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATT 1379
QY 241 TTACTCTCTTTGGAGCCCTGTATCTCACTTCTCTTAAAGTTTGTCTCTCCAGAAAT 300
DB |||||
1380 GTTACTCTCTTTGGAGCCCTGTATCTTAAACCTCTTGTAAAGTTTGTCTCTCCAGAAAT 1439
QY 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAAACCCAGATGCACTGCACTAA 360
DB |||||
1440 TGAAGCTGTAAGCTACAGATGGTCTTACAAATGGAAACCCAGATGCACTGCACTAA 1499
QY 361 AATCTACCGTGAGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
DB |||||
1500 GATCCACCGTGAGACCCCTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTGA 1559
QY 421 AGTCACCCCTCCCGAGGAATCTCACTGCAACAACCCCTACTACACTCCAATTCAGTAGG 480
DB |||||
1560 AGGCACCCCTCCCGAGGAATCTCACTGCAACAACCCCTACTATGCCCAATTCAGCCGG 1619
QY 481 AAGCAGTTAGAGCAGTTGTGAGCCAAACCTCCCAACAGTACTTGGGTTTTCTGTGAGA 540
DB |||||
1620 AAGCAGTTAGAGCGGTCTATCAGCCAACTCCCAACAGTACTTGGGTTTTCTGTGAGA 1679
QY 541 GGGTGGACTGAGAGACAGGACTAGCTGATTTCTAGGCTGACTAAGAAATCCCAAGCT 600
DB |||||
1680 GGGGGGACTGAGAGACAGGACTAGCTGATTTCTAGGCAACGAAAGAAATCCCTAAGCT 1739
QY 601 ANCTGGGAAGGTGACCGCATCTTTTAAACATGGGGCTTGGCACTTAGCTCACACCG 660
DB |||||
1740 AGCTGGGAAGGTGACTGATCCACTCTAAACATGGGGCTTGGCACTTAGCTCACACCG 1799
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC 720
DB |||||
1800 ACCAATCAGAGAGCTCACTAAATGCTAAATAGGCAAAAAATAGGAGTAAAGCAATAGCC 1859
QY 721 AATCATCTATTGCTGAGAGCAGCGGAGGACNAGGATTTGGGATATAAATCACTCAGCA 780
DB |||||
```

Db 1860 AATCATCTATTGCTGAGAGACAGCGGAGGAGCAAGAGATCGGGATATAAAACCCAGGCA 1919
QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCAATGTTATGGAGCTCTGTTTT 840
Db 1920 TTGAGCCGCAACAGCAACCCCTTTGGTCCCTCCCAATGTTATGGAGCTCTGTTTT 1979
QY 841 CACTCTATTCTACTCTATTAAATCATGCAACTGCA 875
Db 1980 CACTCTATTCTACTCTATTAAATCTTGCAACTGAA 2014
RESULT 12
ACN88786
ID ACN88786 standard; DNA; 7974 BP.
XX ACN88786;
XX
XX
XX 02-DEC-2004 (first entry)
XX Breast cancer related marker, seq id 9936.
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX Homo sapiens.
XX US2003099974-A1.
XX 29-MAY-2003.
XX 18-JUL-2002; 2002US-00198846.
XX 18-JUL-2001; 2001US-0306220P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2003-787014/74.
XX
XX Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.
XX Disclosure; SEQ ID NO 9936; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with
XX breast cancer which is encoded by a nucleic acid molecule comprising a
XX nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX the polypeptide of the invention. The activity of the polypeptide of the
XX invention may be described as cytostatic. The antibody is useful for
XX detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX invention are useful in the detection of breast tumours. (I) is useful as
XX a marker for breast cancer and in breast cancer therapy. Sequences given
XX in records ACN78851-ACN92934 represent nucleic acid markers associated
XX with breast cancer. Note: The sequence listing does not form part of the
XX specification but may be obtained in electronic format from the USPTO web
XX site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
XX Sequence 7974 BP; 2034 A; 2017 C; 2123 G; 1800 T; 0 U; 0 Other;
Query Match 58.38; Score 774.2; DB 11; Length 7974;
Best Local Similarity 85.8; Pred. No. 1.2e-242;
Matches 937; Conservative 0; Mismatches 126; Indels 29; Gaps 6;
QY 257 CCCTGTATCTCAACTTCCTTGTAAAGTTGTCTCTCCAGAAATTGAAGCTGTAAGCTA 316
Db 6409 CCCTGTATCTTAACTTCCTTGTAAAGTTGTCTCTCCAGAAATTGAAGCTGTAAGCTA 6468
QY 317 CAAATAGTCTTCAAATGGAAACCCAGATGAGTCCATGAC-TAAATCTACCGTGGACC 375
Db 6469 CAAATCGTCTTCAAATGGAGCCCAAGATGAGTCCATGACTTAAAGATCTACCATGGACC 6528
QY 376 CCTGGACCGGCTGCTAGATGATGCTGATGTTAATGACATTGAAGTACCCCTCCCGA 435

Db 6529 CCTGGACCGGCTGCTAGCCATGCTGATGTTAATGACATCGAAGACACCCCTCCAGA 6588
QY 436 GGAAATCTCAACTGCAACCCCTACTACTACCTCCAAATTCAGTAGGAAGCAGTTAGAGCAG 495
Db 6589 GGAAATCTCAACTGCAACCCCTACTACTATGCCCTGATTCAGCAGGAAGTAGTTAGAGTGG 6648
QY 496 TTGTCAGCCCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGA 555
Db 6649 TTGTCGGCCAACTCCCAACAGCAGTCTGAGTTTCTGTTGGAGGGGGAGCTGAGAGA 6708
QY 556 CAGGACTAGCTGATTTCTCTA-----GGCTGACTTAAGATCCCNAAAGCTTANCTGGGA 608
Db 6709 CAGGACTAGCTGGGATTTCCCTAGGGCCGACTTAAGAATCAACACTAAGCGCTAGCTGGGA 6768
QY 609 AGGTGACCGCAT-CCATCTTTTAAACATGGGGCTTGCAACTTAGCTCACAACCCGACCAATC 667
Db 6769 AGGTGACCGGCTCCCACTTTTAAACACGGGGCTTGCAACTTAGCTCACAACCCGACCAATC 6828
QY 668 -----AGAGAGCTCACTAAATGCTAATCAGSCAAAAAACAGAGGTAAAGCAATAG 718
Db 6829 AGGTAGCAAGAGAGAGCTCGTTAAATGATAATTAGSCAAAAACAGGAGGTAAAGAAATAG 6888
QY 719 CCAATCATC-TAATGCTGAGAGCAGCGGGAAGGACAAGGATTTGGGATATAAACTCAG 777
Db 6889 CCAATCATCTTATCACCTGAGAGCAATATGGGAAGGACAATGATCGGGATATAAAACCCAG 6948
QY 778 GCATTCAGCCAGCAACACCCCTTTGGGTCCTCCCTCCCATCTTATGGGAGCTCTGT 837
Db 6949 GCATTCAGTTGGCAACTCTACCTCTTTGGGTCCCTCCCTTTGTTATGGAAGCTCTGT 7008
QY 838 TTTCACTCTATTCTACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTCTTTA 897
Db 7009 -----TTTCACTCTATTAAATCTTCTGCACTCTTCTGGTCCATATTGTTA 7058
QY 898 TGCTCAAGCTGAGCTTTTGGTCCCATCCACACTGCTGTTGGCCACCGCTCAGACACC 957
Db 7059 CGGCTCCAGCTGAGCTTTTGGTCCCGCTCCACCCCTGTTGTTGGCCGCTGTGCGACACC 7118
QY 958 GCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGCTGCTCTGATCCAGCGA 1017
Db 7119 GCGCTGACTTCCATCCCTCCCGATCCAGAGGTTGTGCTGTGCTCTGATGCGAGCGA 7178
QY 1018 GGTACCCATTGGCAGCTCCCGATCAGGCTAAAGGCTTGCATTTCTCTGATGGCTAAGT 1077
Db 7179 GGGCCCATTTCCACTCTCTGATCGGACTGAGGCTTGCCATTTCTGCAAGCTAAGA 7238
QY 1078 GCCTGGGTTTGTCTTAATAGAACTGAACTGCTGCTCATGGGTTCCATGTTCTTCCAT 1137
Db 7239 GCCTGGGTTTGTCTTAATCGAGCTGAACACTAGTCACTGGGTTCCAGGTTCTTCTCCGT 7298
QY 1138 GACCCAGGCTTCTAATAGAGCTAATCACTACCCATGCGCCCAAGATTCCATTTCTCTTG 1197
Db 7299 GACCCAGGCTTCTAATAGAGCTAATCACTACCCATGCGCCCAAGATTCCATTTCTCTTG 7358
QY 1198 GTATCTGTAGGGCAAGAAACCCAGCTCAGAGAANGTGAAGGCTTGGCCACCATTTGGGAAG 1257
Db 7359 GAATCGTGAAGCAAGAACCCAGCTCAGAGAACACAGGCTTGGCCACCGTCTTGGAG 7418
QY 1258 TGSCCCACTGCCATTTTGGTAGCGGCCCAACCATCTTTGGGAGCTGTGGGAGCAAGAT 1317
Db 7419 TGSCCTGCCACCATTTTGAAGGGGCTGCCACCATCTTTGGGAGCTGTGGGAGCAAGAT 7478
QY 1318 CCCCAGTAACA 1329
Db 7479 CCCCAGTAACA 7490
RESULT 13
AAD41225
ID AAD41225 standard; cDNA; 2074 BP.
XX
XX AAD41225;
AC

XX 30-OCT-2002 (first entry)
 XX Human EMBRY-2 cdna.
 XX
 KW Human; embryogenesis associated protein; AIDS; reproductive disorder;
 KW infertility; endometriosis; endometrial tumour; inflammatory disorder;
 KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;
 KW ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;
 KW EMBRY-2; allergy; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 64..1692
 FT /*tag= a
 FT /product= "EMBRY-2 protein"
 FT sig_peptide 64..123
 FT /*tag= b
 FT mat_peptide 124..1689
 FT /*tag= c
 FT /product= "Mature EMBRY-2 protein"
 XX
 WO200248362-A2.
 XX
 20-JUN-2002.
 XX
 14-NOV-2001; 2001WO-US043956.
 XX
 15-NOV-2000; 2000US-0249407P.
 XX
 (INCY-) INCYTE GENOMICS INC.
 XX
 Ramkumar J, Arvizu C;
 XX
 WPI; 2002-537629/57.
 DR P-PSDB; AAE25054.
 XX
 New polypeptides of human embryogenesis associated proteins for screening
 PT modulators useful for treating or preventing disorders e.g.
 PT endometriosis, infertility, allergy, preeclampsia.
 XX
 Claim 59; Page 96-97; 97pp; English.
 XX
 The invention relates to human embryogenesis associated proteins (EMBR)
 CC and nucleic acid molecules encoding such proteins. EMBRY sequences are
 CC useful for screening modulators useful for treating or preventing
 CC disorders associated with abnormal expression of EMBRY. The disorders
 CC treated include reproductive disorders such as infertility,
 CC endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
 CC disorder such as acquired immune deficiency syndrome (AIDS), allergies,
 CC contact dermatitis; disorders of the placenta such as preeclampsia,
 CC abruptio placentae etc. Sequences of the invention are also useful for
 CC analysing a proteome of a tissue or a cell type. EMBRY proteins are
 CC useful as immunogens for preparing antibodies. Polynucleotides of the
 CC invention are useful for creating knockin humanised animals or transgenic
 CC animals to model human diseases. They are also used in gene therapy. The
 CC present sequence is human EMBRY-2 cdna
 XX
 SQ Sequence 2074 BP; 583 A; 567 C; 392 G; 531 T; 0 U; 1 Other;
 Query Match 57.7%; Score 766.8; DB 6; Length 2074;
 Best Local Similarity 93.0%; Pred. No. 1.4e-240;
 Matches 812; Conservative 0; Mismatches 60; Indels 1; Gaps 1;
 1 TCATAATCGAAGAGCTTTAGACTTCTAACCAGCCAAAAGAGGGGAACTGTTATTTT 60
 1203 TCATAATCGAAGAGCTTTAGACTTCTAACCAGCCAAAAGAGGGGAACTGTTATTTT 1262
 61 AGGGGAAGAAGTCTGTTAGTATGTTAAATCAATCTGGAATCAATTAATGAGAAGTTAAAGA 120
 1263 AGGAGAAGAAGTCTGTTATGTTAAATCAATCAATCCAGAATTTGCTACTGAGAAGTTAAAGA 1322

QY 121 AATTTGAGATCGAATAATAATGTAGAGCAGAGGACCTTCAAAAACACTGCACCCCTGGGGCCT 180
 DB |||||
 QY 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATT 240
 DB |||||
 QY 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTTCTTTAAAGTTTCTCTTCCAGAAAT 300
 DB |||||
 QY 1443 GTTACTCTCTTTGGACCCCTGTATCTTAAACCTCTTTAAAGTTTCTCTTCCAGAAAT 1502
 DB |||||
 QY 301 TGAAGCTGTAAGCTACAATAGTCTTCAATGGAACCCAGATGAGTCCATGACTAA 360
 DB |||||
 QY 1503 TGAAGCTGTAAGCTACAATAGTCTTCAATGGAACCCAGATGAGTCCATGACTAA 1562
 DB |||||
 QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTGTAGACTATGCTCTGATGTTAATGACTTGA 420
 DB |||||
 QY 1563 AATCTACCGGACCCCTGGACCGGCTGTGTAGCCATGCTCCGATGTTAATGACTCGA 1622
 DB |||||
 QY 421 AGTCACCCCTCCGAGGAAATCTCACTGCACACCCCTACTACACTCCAAATTCAGTAGG 480
 DB |||||
 QY 1623 AGGCATCTCTCCCAAGGAAATCTCACTGCACACCCCTACTATGCCCAATTCAGCAGG 1682
 DB |||||
 QY 481 AAGCAGTTAGAGCAGTTGTGTAGCCAACTCTCCCAACAGTACTTGGGTTTTCTCTGTGAGA 540
 DB |||||
 QY 1683 AAGCAGTTAGAGCAGTTGTGTAGTCACTCTCCCAACAGCAGTCTGGGTTTTCTCTGTGAGA 1742
 DB |||||
 QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGCTGACTTAAGAAATCCCNAAAGCCT 600
 DB |||||
 QY 1743 GGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGATTAAGAAATCCCTAAGCCT 1802
 DB |||||
 QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGGCACTTAGCTCACACCCG 660
 DB |||||
 QY 1803 AGCTGGGAAGGTGACCGCTCCACCTTTAAACACGGGGCTTGGCACTTAGCTCACACCCA 1862
 DB |||||
 QY 661 ACCAATCAGAGAGCTCACTAAATGCTAAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC 720
 DB |||||
 QY 1863 ACCAATCAGAGAGCTCACTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCT 1922
 DB |||||
 QY 721 AATCATCTATTGCTGAGAGACAGCGGGAAGGCAAGGATTTGGGATATAAATCAGGCA 780
 DB |||||
 QY 1923 AATCATCTATTGCTGAGAGACAGTGGGGGGAAGGATTTGCAATATAAATCAACCCAGCA 1982
 DB |||||
 QY 781 TTCAAGCCCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
 DB |||||
 QY 1983 TTCAGGCAGC-ANAGCAACCGCTTTGGGTCCCTTCCCTTTGATGGAGCTCTGTTTT 2041
 DB |||||
 QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTG 873
 DB |||||
 QY 2042 CACTCTATTTCACCTCTATTAAATCTTGCAACTG 2074
 DB |||||
 RESULT 14
 ACC46747
 ID ACC46747 standard; cdna; 2046 BP.
 XX
 AC ACC46747;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Human dithp growth/development-associated protein-encoding cDNA.
 XX
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging; growth;
 XX development; gene; ss.
 OS Homo sapiens.

ID AAX77526 standard; cDNA; 2946 BP.
XX
AC AAX77526;
XX
XX
XX 10-AUG-1999 (first entry)
XX
XX Human secreted protein AJ172_2 cDNA.
DE
XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
XX bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
XX cell proliferation; cell differentiation; suppressor; tumour inhibitor;
XX haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
XX haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
XX cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.
XX
XX Homo sapiens.
XX
XX WO9926972-A1.
XX
XX 03-JUN-1999.
XX
XX 17-NOV-1998; 98WO-US024614.
XX
XX 21-NOV-1997; 97US-00976110.
XX
XX 18-MAY-1998; 98US-00080478.
XX
XX 20-OCT-1998; 98US-00175928.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Treacy M;
XX
XX WPI; 1999-357813/30.
XX
XX P-PSDB; AAY08622.
XX
XX New polynucleotides encoding secreted proteins.
XX
XX Claim 13a; Page 100-101; 142pp; English.
XX
XX This invention describes novel human secreted proteins encoded by
XX polynucleotides isolated from human adult testes, adult brain, adult
XX blood or adult placenta, or murine adult bone marrow or thymus cDNA
XX libraries. The products of the invention are predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals, although no
XX supporting data is given. Suggested activities include nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immune stimulating (e.g. as vaccines) or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, and tumour
XX inhibition activity. The polynucleotides are also stated to be useful for
XX gene therapy
XX
XX Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;
SQ
Query Match 55.1%; Score 731.8; DB 2; Length 2946;
Best Local Similarity 90.7%; Pred. No. 5.7e-229;
Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
QY 1 TCANAATCGAAGCTTTAGACTTCTACCGCCCAAGAGGGGGAACCTGTTATTTT 60
DB 2067 TCANAATCGAAGCTTTAGACTTCTACCGCCCTGAAAGAGGGGGAACCTGTTATTTT 2126
QY 61 AGGGGAAGAATCCTGTTAGTATGTTAATCAATCTGGAATCATTACTGAGAAGTTAAAGA 120
DB 2127 AGGGGAAGNATGCTGTTATGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAAGA 2186
QY 121 AATTTGAGATCGAATATATATGATAGACGAGACCTTCAAAACACTGCACCTGGGGCT 180
DB 2187 AATTCGAGATCGAATACACGATAGACGAGGAGCTTCGAAACACTGGACCTGGGGCT 2246

QY 181 CCTCAGCCAAATGGATGCCCTGGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
DB 2247 CCTCAGCCAAATGGATGCCCTGGATTTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 2306
QY 241 TTTACTCTCTTTTGGACCCCTGTATCTTCAACTTCTCTTAAAGTTTGTCTCTCCAGAAT 300
DB 2307 GCTACTCTCTTTGGACCCCTGTATCTTAACTCTCTTGTAACTTGTCTCTCCAGAAT 2366
QY 301 TGAAGCTGTAAGCTACAATAGTTCTTCAATAGGAACCCAGATGAGTGCATGACTAA 360
DB 2367 CGAAGCTGTAAAACCTA-----CAAATGGAGCCCAAGATGCAATGCAAGACTAA 2414
QY 361 AATCTACGTTGGACCCCTGTCGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 2415 GATCTACCGAGACCCCTTGACCGGCTGTAGCCGAGATCTGATGTTAATGACATCAA 2474
QY 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCAACAAACCCCTACTACATCTCAATTCAGTAGG 480
DB 2475 AGGCACCCCTCCTGAGGAAATCTCAGCTGCAACACCTCTACTAGCCCCCAATTTCAGCAGG 2534
QY 481 AAGCAGTTAGAGCAGTTGTGAGCCAACTCCGCCAAGTACTTGGGTTTCTCTGTGAGA 540
DB 2535 AAGCAGTTAGAGCAGTCTGTCGGCCAACTCCGCCAAGTACTTGGGTTTCTCTGTGAGA 2594
QY 541 GGGTGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCNAAAGCCT 600
DB 2595 TGGGGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCNAAAGCCT 2654
QY 601 ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGCACCTTAGCTCACACCGG 660
DB 2655 AGCTGGGAAGGTGACCATCTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2714
QY 661 ACCAATCAGAGCTCACTAAATGCTAATCAGCGCAAAACAGGAGTAAAGCAATAGCC 720
DB 2715 ACCAATCAGAGCTCACTAAATGCTAATCAGCGCAAAACAGGAGTAAAGCAATAGCC 2774
QY 721 AATCATCTATTGCTGAGACGACAGCGGGAAGCAAGGATTGGGATATAAATCTCAGGCA 780
DB 2775 AATCATCTATTGCTGAGACGACAGCGGGAAGCAAGGATTGGGATATAAATCTCAGGCA 2834
QY 781 TTCAAGCCGACCAAGCAACCCCTTTGGGTCCTCCCTCCATTTGATGGGAGCTCTGTTTT 840
DB 2835 TTCAGCGCGCAAGCGCAACCCCTTTGGGTCCTCCCTCCATTTGATGGGAGCTCTGTTTT 2894
QY 841 CACTCTATTTTCACTCTATTAAATCATCAACTGCA 875
DB 2895 CATGCTATTTTCACTCTATTAAATCTTGCAACTGCA 2929
RESULT 16
AAZ59468
ID AAZ59468 standard; cDNA; 2946 BP.
XX
AC AAZ59468;
XX
XX 11-APR-2000 (first entry)
DT
XX Human secreted protein AJ172_2 polynucleotide sequence.
DE
XX Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
XX placental pathology; metastasis inhibition; nutritional activity;
XX immune stimulator; haematopoiesis regulator; tissue growth;
XX tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
XX gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO9960020-A1.
XX
XX 25-NOV-1999.
XX
XX 17-MAY-1999; 99WO-US010915.
XX
XX

PR 18-MAY-1998: 98US-00080478.
 PR 20-OCT-1998: 98US-00175928.
 XX (GEMY) GENETICS INST INC.
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Mi S, Treacy M;
 XX WPI; 2000-116311/10.
 DR P-PSDB; AAY67313.
 XX
 XX New polynucleotides encoding secreted cDNA libraries, used to develop
 PT products for the diagnosis and treatment of neoplastic disease.
 XX
 XX Claim 14; Page 107-108; 149pp; English.
 XX
 CC This is the human secreted protein AJ172.2 nucleotide sequence, obtained
 CC from a human adult testes cDNA library. The invention relates to secreted
 CC human and murine proteins. The polynucleotides and proteins are predicted
 CC to have biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals. Detection of the levels of the proteins can be used for the
 CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
 CC which modulate the expression or function of the proteins may be used for
 CC treating a neoplastic disease and inhibiting metastasis. Other suggested
 CC activities include nutritional activity (e.g. in feeds), cytokine and
 CC cell proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity. The
 CC polynucleotide sequences are also stated to be useful for gene therapy
 XX
 XX Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;
 SQ

Query Match 55.1%; Score 731.8; DB 3; Length 2946;
 Best Local Similarity 90.7%; Pred. No. 5.7e-229;
 Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGAACTGTTTATTTT 60
 2067 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGAACTGTTTATTTT 2126
 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCAATCTGAGAACTTAAAGA 120
 2127 AGGGGAAGATGCTGTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 2186
 121 AATTTGAGATCGAATATAATGTAGACGAGGAGCTTCAAAACACTGCACCTGGGGCCT 180
 2187 AATTCGAGATCGAATACAGCTAGACGAGGAGCTTCGAAACACTGGACCTGGGGCCT 2246
 181 CCTCAGCAATGAGTCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 240
 2247 CCTCAGCAATGAGTCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 2306
 241 TTTACTCTCTTTGGACCTGTTATCTCAACTCTCTTTTAACTTTGTTCTTCCAGAT 300
 2307 GCTACTCTCTTTGGACCTGTTATCTTAACTCTCTTTTAACTTTGTTCTTCCAGAT 2366
 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGAGTCCAGACTAA 360
 2367 CGAAGCTGTAAGCTA-----CAATGGAAGCCCAAGATGAGTCCAGACTAA 2414
 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTGATGTTTAAATGACATGTA 420
 2415 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 2474
 421 AGTCACCCCTCCCGAGGAATCTCAACTGACAAACCCCTACTACTCCAAATTCAGTAGG 480
 2475 AGGCACCCCTCTCCTGAGGAATCTCAGCTGACAAACCTCTACTAGCCGCCAAATTCAGCAGG 2534
 481 AAGCAGTTAGCAGGTTGTCAGCAACCTCCCAACAGTACTTTGGGTTTCTCTGTTGAGA 540

Db 2535 AAGCAGTTAGACGGTCTGCGCCAACTCCCAACAGCAGCTAGCTTTTCTGTTGAGA 2594
 QY 541 GGGTGGACTGAGACAGAGCTAGCTGGATTTCTTAGCTAGCTAAGAAATCCNAGCCT 600
 Db 2595 TGGGGGACTGAGACAGAGCTAGCTGGATTTCTTAGCTAGCTAAGAAATCCNAGCCT 2654
 QY 601 ANCTGGGAAGGTGACCGCATCTTTTAAACATGGGGCTTCAACTTAGCTACACCCG 660
 Db 2655 AGCTGGGAAGGTGACCACTCCACCTTTAAACACGGGGCTTCAACTTAGCTACACCTG 2714
 QY 661 ACCAATCAGAGAGCTACATAAATCTATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
 Db 2715 ACCAATCAGAGAGCTACATAAATCTATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2774
 QY 721 AATCATCTATTGCTGAGACAGCGGAGGACAGCAAGGATTTGGGATATAAACTCAGGCA 780
 Db 2775 AATCATCTATTGCTGAGACAGCGGAGGACAGCAAGGATTTGGGATATAAACTCAGGCA 2834
 QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCACTTGTATGGGAGCTCTGTTT 840
 Db 2835 TTGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCACTTGTATGGGAGCTCTGTTT 2894
 QY 841 CACTCTATTCTCTCTATTAAATCATGCAACTGCA 875
 Db 2895 CATGCTATTCTCTCTATTAAATCTTGCAACTGCA 2929

RESULT 17
 ADC38776
 ID ADC38776 standard; cDNA; 2946 BP.
 XX
 AC ADC38776;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human cDNA encoding a secreted protein #63.
 XX
 KW ss: gene; immune disorder; severe combined immunodeficiency; SCID;
 KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
 KW rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;
 KW lymphoid cell deficiency; osteoporosis; osteoarthritis;
 KW peripheral nervous system disease; peripheral neuropathy;
 KW Alzheimer's disease; Parkinson's disease; coagulation disorder;
 KW inflammatory disease; systemic inflammatory response syndrome; SIRS;
 KW ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;
 KW hypersensitivity; regeneration; neural cell proliferation; fertility;
 KW tumour; chemokine; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 XX US2002193567-A1.
 PN
 XX 19-DEC-2002.
 PD
 XX
 PF 02-APR-2002; 2002US-00114893.
 XX
 XX 11-AUG-1995; 95US-00514014.
 PR 05-APR-1996; 96US-00628364.
 PR 19-APR-1996; 96US-00635311.
 PR 07-JUN-1996; 96US-00659224.
 PR 17-JUN-1996; 96US-00664596.
 PR 09-JUL-1996; 96US-00677231.
 PR 26-JUL-1996; 96US-00686878.
 PR 23-AUG-1996; 96US-00701819.
 PR 27-SEP-1996; 96US-00721488.
 PR 27-SEP-1996; 96US-00721798.
 PR 27-SEP-1996; 96US-00721923.
 PR 27-SEP-1996; 96US-00721926.
 PR 25-OCT-1996; 96US-00738367.
 PR 30-OCT-1996; 96US-00739775.
 PR 13-JAN-1997; 97US-00783395.
 PR 10-APR-1997; 97US-00833823.

[illegible]

PT developing preclampsia, comprises determining whether the expression or
PT activity of syncytin in the cell is modulated in the presence of a test
XX compound.

XX Disclosure; Page 39-42; 43pp; English.

PS The invention relates to identifying compounds which are modulators of
XX syncytin expression. The syncytin modulators are useful in diagnosis and
XX treatment of preclampsia and gestational trophoblast disorders (e.g.
XX choriocarcinoma, hydatiform mole, placental site tumour and missed/
XX incomplete abortion). Syncytin is a human gene derived from the envelope
XX gene of human endogenous defective retrovirus, HERV-W. The present
XX invention is based partly on the discovery that syncytin expression is
XX dramatically reduced in preclampsia, and is also mis-localised to the
XX apical syncytiotrophoblast membrane. The present sequence is human
XX syncytin cDNA

SQ Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;

Query Match 55.0%; Score 730.8; DB 6; Length 2930;
Best Local Similarity 90.7%; Pred. No. 1.2e-228;
Matches 793; Conservative 0; Mismatches 69; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCMAAGAGGGGAACTGTTTATTTT 60
DB TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGGGGAACTGTTTATTTT 2128
QY 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
DB AGGGGAAGAAATGCTGTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 2188
QY 121 AATTTGAGATCGAATATATATAGACAGAGGACCTTCAAACACACTGCACTTGGGCGCT 180
DB AATTCGAGATCGAATATCAACAGTAGACAGAGGAGCTTCCGAAACACTGGACCTTGGGCGCT 2248
QY 181 CCTCAGCCAAATGATCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
DB CCTCAGCCAAATGATGCCCTGGATTTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 2308
QY 241 TTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTTGTAACTTTGCTCTCTTCCAGAA 300
DB GCTACTCTCTTTGGACCCCTGATCTTCAACTCTCTTGTAACTTTGCTCTCTTCCAGAA 2368
QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAAACCCAGACAGTCCAGTAA 360
DB CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGAGTCCAGGACTAA 2416
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
DB GATCTACCGGAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 2476
QY 421 AGTCACCCCTCCGAGGAAATCTCAATGCACAAACCCCTACTACATCCCAATTCAGTAGG 480
DB AGGCACCCCTCCGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 2536
QY 481 AAGCAGTTAGACAGTTGTCAGGCAACCTCCCAACAGTACTTGGGTTTCCCTGTGAGA 540
DB AAGCAGTTAGAGGGGTGCTGGGCAACCTCCCAACAGCACTTAGGTTTCTCTGTGAGA 2596
QY 541 GGGTGGACTCAGAGACAGGACTAGTGGATTCTCTAGGCTGACTAAGAAATCCCNAAAGCCT 600
DB TGGGGACTCAGAGACAGGACTAGTGGATTCTCTAGGCTGACTAAGAAATCCCTAGCCT 2656
QY 601 ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGAACCTTAGCTCACACCCG 660
DB AGCTGGGAAGGTGAACACATCCACCTTTAAACACGGGGCTTGAACCTTAGCTCACACCTG 2716
QY 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACACAGAGTAAGCAATAGCC 720
DB ACCAATCAGAGAGCTCACTAAATGCTTAATTAAGCAAAACACAGAGTAAGCAATAGCC 2776
QY 721 AATCATCTATTGCTGAGACACAGCGGGAGGACAGGATTTGGGATATATAAATCTCAGGCA 780

DB 2777 AATCATCTATTGCTGAGAGCACAGAGGAGCAATGATCGGATATAAACCAAGTC 2836
QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTT 840
DB 2837 TTCGAGCCGCAACGCAACCCCTTTGGTCCCTCCCTTTGATGGAGCTCTGTTT 2896
QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGC 874
DB 2897 CATGCTATTTCACCTCTATTAAATCTTGCAACTGC 2930

RESULT 19
AAF55630
ID AAF55630 standard; DNA; 2781 BP.
XX AAF55630;
XX
XX 29-MAY-2001 (first entry)
XX Nucleotide sequence of a human endogenous retrovirus envelope protein.
XX Envelope protein; HERV; syncytia formation; placental development;
XX syncytia; cancer; cell adhesion; ss.
XX Human endogenous retrovirus.
XX
XX Key Location/Qualifiers
XX CDS 762..2378
XX FT /*tag= a
XX FT /product= "envelope protein"
XX
XX WO200116171-A1.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-FR002429.
XX
XX 01-SEP-1999; 99FR-00011141.
XX 15-SEP-1999; 99FR-00011793.
XX
XX (INMR) BIO MERIEUX.
XX (INRM) INST NAT SANTE & RECH MEDICALE.
XX
XX Mallet F, Cosset F, Blond J, Lavillette D, Bouton O, Ruggieri A;
XX WPI; 2001-226676/23.
XX P-PSDB; AAB67652.
XX
XX Detecting expression of human endogenous retrovirus envelope protein in
XX cells of a tissue or culture, from its ability to induce syncytia.
XX
XX Disclosure; Page 44-45; 57pp; French.
XX
XX The present sequence encodes a human endogenous retrovirus envelope
XX protein. The specification describes a method for detecting expression of
XX an envelope protein from a human endogenous retrovirus (HERV), in cells,
XX of a tissue or culture. The method comprises detecting syncytia formation
XX due to the fusogenic properties of the envelope protein. Envelope
XX polypeptides and polynucleotides are used to produce therapeutic or
XX prophylactic compositions, particularly for treatment of cancer, to
XX correct defects in placental development (or other natural formation of
XX other types of syncytia), and to promote adhesion of cells in grafts or
XX cellular repair processes. Expression of sequences antisense to the
XX polynucleotide are used to prevent formation of syncytia
XX
XX
XX Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 U; 0 Other;

Query Match 54.9%; Score 730.2; DB 5; Length 2781;
Best Local Similarity 90.6%; Pred. No. 1.8e-228;
Matches 793; Conservative 0; Mismatches 70; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTTTATTTT 60

Db 1901 TCAAAATCGAAGAGCTTTAGACTTGCCTAAACCGCTGAAGAGAGGGGAAACCTGTTATATTTT 1960
Qy 61 AGGGAAGAATGCTGTTAGTATGTTTAAATCAATCTGGAAATCAATCTGAGAAGTTAAAGA 120
Pd 1961 AGGGAAGAATGCTGTTATGTTTAAATCAATCGGAATCGTCACTGAGAAGTTAAAGA 2020
Qy 121 AATTGAGATCGAATATATATAGTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 2021 AATTGAGATCGAATATCAACGCTAGAGCAGAGGAGCTTCGNAACACTGGACCTGGGGCT 2080
Qy 181 CCTCAGCAATGAGTCCCTGACCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 2081 CCTCAGCAATGAGTCCCTGAGTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 2140
Qy 241 TTTACTCTCTTTTGGACCTGATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCAGAAAT 300
Db 2141 GCTACTCTCTTTGGACCTGATCTTAAACCTCTTGTAACTTGTCTCTTCAGAAAT 2200
Qy 301 TGAAGCTGAAGCTACAATAGTTCTTCAATGGAACCCAGATGCGAGTCCATGACTAA 360
Db 2201 CGAAGCTGAAGCTA-----CAAATGAGGCCAAGATGCGAGTCCAAGACTAA 2248
Qy 361 AATCTACCGTGGACCCCTGGAGCCGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2249 GATCTACCGCAGACCCCTGGAGCCGCTGTAGCCACGATCTGATGTTAATGACATCA 2308
Qy 421 AGTACCCCTCCGAGGAAATCTCAACTGACAAACCCCTACTACACTCCAATTCAGTAGG 480
Db 2309 AGGACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 2368
Qy 481 AAGCAGTTAGACAGTCTGACGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db 2369 AAGCAGTTAGACGCTGCTGCGCCAACTTCCCAACAGACTTGGTTTCTCTGTTGAGA 2428
Qy 541 GGGTGGACTGAGAGCAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 2429 TGGGGACTGAGAGCAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT 2488
Qy 601 ANCTGGGAGTGAACCGATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 2489 AGCTGGGAGTGAACCATCCACTTTAAACACAGGGCTTTGCAACTTAGCTCACACCTG 2548
Qy 661 ACCAATCAGAGCTCCTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2549 ACCAATCAGAGCTCCTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCC 2608
Qy 721 AATCATCTATTGCTGAGAGCACAGCGGAGGCAAGGATTTGGATATAAATCAGGCA 780
Db 2609 AATCATCTATTGCTGAGAGCACAGCAGGAGGCAATGATCGGGATATAAACCAGTC 2668
Qy 781 TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTT 840
Db 2669 TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTT 2728
Qy 841 CACTCTATTCTACTTATTAAATCATGCAACTGCA 875
Db 2729 CATGCTATTCTACTTATTAAATCTTGCAACTGCA 2763

RESULT 20
ID AAS84210 standard; cDNA; 6394 BP.
XX AC AAS84210;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #20014.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.

XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG20023.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 20014; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridization probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 6394 BP; 1840 A; 1597 C; 1384 G; 1571 T; 0 U; 2 Other;
Query Match 54.8%; Score 728.6; DB 5; Length 6394;
Best Local Similarity 78.6%; Pred. No. 1e-227;
Matches 1094; Conservative 0; Mismatches 223; Indels 74; Gaps 16;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGCCTAAACCGCTGAAGAGAGGGGAAACCTGTTATTTT 60
Db 1175 TCAAAATCGAAGAGCTTTAGACTTGCCTAAACCGCTGAAGAGAGGGGAAACCTGTTATTTT 1234
Qy 61 AGGGAAGAATGCTGTTAGTATGTTTAAATCAATCTGGAAATCAATCTGAGAAGTTAAAGA 120
Db 1235 AGGGAAGAATGCTGTTATGTTTAAATCAATCGGAATCGTCACTGAGAAGTTAAAGA 1294
Qy 121 AATTGAGATCGAATATATAGTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 1295 AATTGAGATCGAATATCAACGCTAGAGCAGAGGAGCTTCGNAACACTGGACCTGGGGCT 1354
Qy 181 CCTCAGCAATGAGTCCCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 1355 CCTCAGCAATGAGTCCCTGAGTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 1414
Qy 241 TTTACTCTCTTTTGGACCTGATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCAGAAAT 300
Db 1415 GCTACTCTCTTTGGACCTGATCTTAACTCTCTTGTAACTTTGTTCTCTTCAGAAAT 1474
Qy 301 TGAAGCTGAAGCTACAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCAAGACTAA 360

1475 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCACTCCAGACTAA 1522
361 AATCTACCGTGGACCCCT-GGACCGGCTGCTAGACTATGCTCTGATGTATATGACATTTG 419
1523 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTATATGACATCA 1582
420 AAGTACCCCTCCGAGGAATCTCACTGACGCAACCCCTACTACACTCC-----469
1583 AAGGCACCCCTCTGAGGAAATCTAGCTGCAAAAACCTCTACTTACGCCCCCAATTCAGG 1642
470 -----AATTCACTAGGAAGCAGTTAGAGCAGTGTGTGAGCCAACTCCCAACAGTACTTTG 524
1643 CAGGGAAGCAGTTAGAGCGCGGTCTGTCGGGCCAACTTCCCAACAGGCACTTTAGG 1702
525 GGTTCCTGTTGAGAGGTGGACTGAGAG-----ACAGACTAGCTTGGGA--TTTCCTAGG 578
1703 GTTTTCTGTTGGAGATGGGGGACTGAGGAGACAGGAGATTAGCTGGGATTTCTCTTAGG 1762
579 CTGACTAAGNATCCCNAGCCTANCTGGGNAAGTGACCG-----CATCCATCT 626
1763 CTGACCTAAGATCCCTTAAGNCTTAGCTGGGGAAGGTGGACCAATCCACCTTTT 1822
627 TTAACATGGG---CTTGCACTTAGCTCACACCC---GACCAATCAGAGAGCTCACTA 680
1823 TAAACACGGGGGCTTTGCACTTTAGTTTACACTTGCACCAATCAGAGAGCTCACTA 1882
681 AATG-CTAATCAGGCAAAACAGGAGTTAAGCAATAGCAATCATCTATTGCTCTGA-G 738
1883 AATGCTTAATTAGGCAACAGAGAGGTAAGAAATAGCAATCATCTATTGCTCTGAGG 1942
739 AGCAGAGCGGAAGGACAGGATTGGGATATAACTCAGGCATTCAAGCCAGC-RACAGC 797
1943 AGCAGAGTGGGAGGCAAGGATTGCAATATAAACCCAGGCACTTCGAGCCAGGAGCGC 2002
798 AACCCCTTTGGGTCCCTCCCACTGATGAGGAGCTCTGT-----TTTCACTCTA 847
2003 AACCGCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCATGCTATTTCACCTCTA 2062
848 TTTCACTCTATAAATCATGCAACTG--CACTCTTCTGTCGCTGTTTATATGCTCTCAA 905
2063 TTTCACTCTATAAATCTTGCAACTGTACACTCTTCTGTCAGTGTATTTTACAGCTCTA 2122
906 GCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGGCACCCGCA---CAGACCCGCTGC 962
2123 GCTGAGCTTCTGTTGCACTCCACCTGCTGTTTGGCGCGCCACCGCAGACCCGCTGC 2182
963 TGACTTCCATCCCTTTGGATCCAGAGAGTGTCCACTGTCTCTGATCCAGCGAGGTAC 1022
2183 TGGCTTCCATCCCTTCCGATCCGCGAGGTTGTCTGTCGATTCCTGATCCAGCGAGCGC 2242
1023 CCATTGCACTCCCGATCAGGCTAAAGGCTTGCCATTTGTCGATGCTAGTGGCTG 1082
2243 CCATTGCAACGCCCAATCGGCTAAAGGCTTGCCATTTGTCGATGCTAGTGGCTG 2302
1083 GGTTCGCTTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142
2303 GGTTCGCTTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2362
1143 ACGCTTCTTAATAG--AGCTATTAACACTCACCGCATGGCCCAAGATTCCTTCTGCTGA 1200
2363 ATGCTTCTTAATAGAACTATTAACTATCAACATGATGGCCCAAGATTCCTTCTGCTGA 2422
1201 --TCTGTAGGCCAAGAACCCAGGTGAGAGANGTGGAGTTGCTGCAACCAATTTGGGAGT 1258
2423 ATTCGCTGAGGCCAAGAACCCAGGTGAGAGANGTGGAGTTGCTGCAACCAATTTGGGAGT 2482
1259 GGCCCACTGCCATTTGCTAGCGGCCCAACCACTTTGGAGCTGTTGGGAGCAAGGATC 1318
2483 GGCCCGTGGCCATCTGGAAGCGGCTTGGCCCAACCACTTTGGAGCTGTTGGGAGCAAGGATC 2542
1319 CCCAGGTAACA 1329
|||||

Db 2543 CCCCGGTAAACA 2553
RESULT 21
AA25565
ID AAX25665 standard; cDNA to mRNA; 7582 BP.
XX
AC AAX25665;
XX
DT 21-MAY-1999 (first entry)
XX
DE Complete human endogenous retrovirus W genome.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.
XX
PF 06-JUL-1998; 98WO-FR001442.
XX
PR 07-JUL-1997; 97PR-00008815.
XX
PA (INMR) BIO MERIEUX.
XX
PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX
PS WPI; 1999-120897/10.
XX
PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.
XX
PS Claim 1; Page 71-74; 106pp; French.
XX
CC This sequence represents the complete sequence of the human endogenous
CC retrovirus (HERV) W genome. The nucleic acids, their fragments or
CC peptides encoded by them are markers of autoimmune disease (e.g. multiple
CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
CC insulin- dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility
XX
SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;
Query Match 54.4%; Score 722.6; DB 2; Length 7582;
Best Local Similarity 88.5%; Pred. No. 1.1e-225;
Matches 774; Conservative 19; Mismatches 70; Indels 12; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGCTAACCCGCAAAAGAGGGGAACTGTTATTTT 60
6720 TCAAAATCGAAGAGCTTTAGACTTGCTAACCCGCTGAGAGAGGGGAACTGTTATTTT 6779
QY 61 AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
6780 AGGGGAAGAAATGCTGTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 6839
QY 121 AATTTCAGATCGAATATAATGATAGCAGAGGACCTTCAAAACACTGCACCCCTGGGCGCT 180
6840 AATTTCAGATCGAATATAACACTGAKAGAGAGCTTCGAAACACTGGACCCCTGGGCGCT 6899
QY 181 CCTCAGCCAAATGGATGCCCTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
6900 CCTCAGCCAAATGGATGCCCTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 6959
QY 241 TTTACTCTCTTTGGAGCCCTGATCTTCAACTTCTTGTAGTTGCTCTTCCCAAGAT 300
6960 GCTACTCTCTTTGGAGCCCTGATCTTCACTCTTGTACCTTGTCTTCTTCCCAAGAT 7019
Db

QY	301	TGAAGCTGTAAGAGCTACAAATAGTTCCTTCAAATGGAAACCCAGATGCAGTCCATGACTAA	360
Db	7020	CGAAGCTGTGTRAACTA-----CAAATGGAGCCCAAGATGCAGTCCAAAGACTAA	7067
QY	361	AATCTACCGTGGACCCCTGGACCGGCTGTCTAGACATATGCTCTGATGTTTAATGACATTTGA	420
Db	7068	GATCTACCGCAGACCCCTGGACCGGCTGTCTAGACCCAGCATCTGATGTTTAATGACATCAA	7127
QY	421	AGTCACCCCTCCCGAGGAATCTCACTGACAAACCCCTACTACACTCCCAATTCAGTACG	480
Db	7128	AGGCACCCCTCCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG	7187
QY	481	AAGCAGTTAGACGAGTTGTCTAGCCCAACCTCCCAACAGTACTTGGGTTTTCTCTGTTGAGA	540
Db	7188	AAGCAGTTAGACGCTGCTGGCCAACTCCCAACAGCAGTCTAGGTTTCTCTGTTGAGA	7247
QY	541	GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGCTGACTTAAGAATCCCNAGCCT	600
Db	7248	TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGCTGACTTAAGAATCCYTAAGCCT	7307
QY	601	ANCTGGGAAGGTGACCCGATCCATCTTTAAACATGCGGCTTGCACACTTAGCTCACACCCG	660
Db	7308	AGSTGGGAAGGTGACCAATCCATCTTTAAACAGCGGGCTTGCACACTTAGCTCACACCTG	7367
QY	661	ACCAATCAGAGAGCTCACTAAATGCTAAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	720
Db	7368	ACCAATCAGAGAGCTCACTAAATGCTAAATGAGCAAGAGGAGTAAAGAAATAGCC	7427
QY	721	AATCATCTATTTCCTGAGAGACAGCGGGAAGGACAAAGGATTTGGGATATAAATCAGGCA	780
Db	7428	AATCATCTATTTCCTGAGAGACAGCGGGAAGGACAAATGATTCGGGATATAAATCAGGCA	7487
QY	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTATGATCGGAGCTCTGTTTT	840
Db	7488	TTCAAGCCGGAACAGGCAACCCCTTTGGGTCCCTCCCTATGATCGGAGCTCTGTTTT	7547
QY	841	CACCTCTATTTCACCTCTATTAAATCATGCAACTGCA	875
Db	7548	CATGCTATTTCACCTCTATTAAATCATGCAACTGCA	7582
RESULT 22			
ID	AAA59215	standard; DNA; 7582 BP.	
XX	AAA59215;		
AC	AAA59215;		
DT	07-NOV-2000	(first entry)	
XX	Human endogenous retrovirus W (HERV-W) sequence.		
DE	Human endogenous retrovirus; human endogenous retrovirus W; HERV-W;		
KW	gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.		
XX	Human endogenous retrovirus.		
XX	Human endogenous retrovirus.		
FH	Key	Location/Qualifiers	
FT	LTR	1..120	
FT		/tag= a	
FT		/note= "R of 5' LTR"	
FT		121..575	
FT		/tag= b	
FT		/note= "U5 of 5' LTR"	
FT	primer_bind	579..596	
FT		/tag= c	
FT		581..7194	
FT	CDS	/tag= d	
FT		/note= "ORF1 env538"	
FT		7039..7194	
FT		/tag= e	
FT		/note= "ORF2 52 AA"	
FT		7112..7255	
FT	CDS		

FT		/tag= f	
FT	misc_feature	/note= "ORF3 48 AA"	
FT		7244..7254	
FT		/tag= g	
FT		/note= "polypurine tract"	
LTR		7256..7582	
FT		/tag= h	
FT		/note= "U3-R of 3' LTR"	
FT	polyA_signal	7563..7569	
FT		/tag= i	
PN	WO20043521-A2.		
XX			
XX	27-JUL-2000.		
PD			
XX			
XX	21-JAN-2000; 2000WO-FR000144.		
Pf			
XX			
XX	21-JAN-1999; 99FR-00000888.		
PR			
XX	(INMR) BIO MERIEUX.		
PA			
XX	Paranhos-Baccala G, Mallet F, Voisset C;		
PI			
XX			
XX	WPI; 2000-499229/44.		
DR			
XX			
PT	New nucleic acid from human endogenous retrovirus, useful e.g. for		
XX	diagnosis of autoimmune disease and complications of pregnancy, contains		
PT	at least part of the gag gene.		
PT			
XX			
XX	Disclosure; Page 49-52; 53pp; French.		
PS			
XX			
CC	The present sequence represents an endogenetic retrovirus, which is		
CC	associated with an autoimmune disease, and is integrated into the human		
CC	genome. The retrovirus is human endogenous retrovirus W (HERV-W). The		
CC	HERV-W retrovirus is associated with autoimmune disease, failure of		
CC	pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or		
CC	proteins derived from it, are useful for diagnosis of autoimmune disease		
CC	(specifically multiple sclerosis) and for monitoring pregnancy. The		
CC	nucleic acid fragments may also be used for in situ labelling of isolated		
CC	chromosomes, while the transcription product can be used to study or		
CC	monitor T cell proliferation in vitro		
XX			
XX	Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;		
SQ			
	Query Match	54.4%; Score 722.6; DB 3; Length 7582;	
	Best Local Similarity	88.5%; Pred. No. 1.1e-225;	
	Matches	774; Conservative 19; Mismatches 70; Indels 12; Gaps 1;	
Qy	1 TCRAATCGAAGACTTTTAGACTTGTCTAACCGCAAAAGAGGGGGAACCTGTTATTATTTT	60	
Db	6720 TCRAATCGAAGACTTTTAGACTYGTCTAACCGCTGARAGGGGGAACCTGTTATTATTTT	6779	
Qy	61 AGGGGAAGAATGCTGTTAGTATGTTTAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA	120	
Db	6780 AGGGGAAGAATGCTGTTATTTATCTTAATCAATCCGAATCGTCACTGAGAAAGTTAAAGA	6839	
Qy	121 AATTGAGATCGAATATATATGTAGACAGAGGACCTTCAAACAACACTGCACCTGGGGCCT	180	
Db	6840 AATTCGATCGAATACACGATKAGCAGARGAGCTTCGNAACACTGGACCTGGGGCCT	6899	
Qy	181 CCTCAGCCAATGGATGCCCTGGACTCTCCCTCTTTAGGACCTTAGCAGCTATAATATT	240	
Db	6900 CCTCAGCCRATGGATGCCCTGGATTCTCCCTCTTTAGGACCTTAGCAGCTATAATATT	6959	
Qy	241 TTACTCTCTTTGGACCCTGTATCTCAACTTCCTTGTTAAGTTGTCTCTCTCCAGAAT	300	
Db	6960 GCTACTCTCTTTGGACCCTGTATCTTTTACCTCCCTTGTTAACCTTTGTCTCTCTCCAGAAT	7019	
Qy	301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAA	360	
Db	7020 CGAAGCTGTAACTA-----CAAATGGAGCCCCAAGATGCAGTCCCAAGACTAA	7067	
Qy	361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGTATGTTAATGACATTGA	420	

||||| 7068 GATCTACCGACGACCCCTCGACCGGCTGTAGCCACCATCTGATTTAATGACATCAA 7127
Db
QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATCTAGTAGG 480
Db 7128 AGGCACCCCTCTCGAGGAAATCTCAGCTGCACAAACCTCTACTACGCGCCCAATTCAGCAGG 7187
QY 481 AAGCAGTTAGACAGTTGTTCAGCCACCTCCCAACAGTACTTGGCTTTCTGTTGAGA 540
Db 7188 AAGCAGTTAGACAGTTGTTCAGCCACCTCCCAACAGTACTTGGCTTTCTGTTGAGA 7247
QY 541 GGGTGGACTCAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT 600
Db 7248 TGGGGACTCAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT 7307
QY 601 ANCTGGGAAGGTGACCGCATCTCAATCTTTAAACATGCGGGCTTGCACCTAGCTCAGACCCG 660
Db 7308 AGSTGGGAAGGTGACCGCATCTCAATCTTTAAACATGCGGGCTTGCACCTAGCTCAGACCCG 7367
QY 661 ACCAATCAGAGGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 7368 ACCAATCAGAGGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 7427
QY 721 AATCATCTATTGCTCAGAGCAGACGCGGAAGGACAAAGGATTGGGATATAAACTCAGGCA 780
Db 7428 AATCATCTATTGCTCAGAGCAGACGCGGAGGACAAATGATCGGATATAAAACCAAGTY 7487
QY 781 TTCAAGCCAGCAACAGAAACCCCTTTGGGTCCCTCCCAATGATGGAGCTCTGTTTT 840
Db 7488 TTCAAGCCAGCAACAGAAACCCCTTTGGGTCCCTCCCAATGATGGAGCTCTGTTTT 7547
QY 841 CACTCTATTTCACCTATTAAATCATGCAACTGCA 875
Db 7548 CATGCTATTTCACCTATTAAATCTTGCACCTGC 7582

RESULT 23

AA25660
ID AAX25660 standard; cDNA to mRNA; 1136 BP.
AC AAX25660;
XX
XX 21-MAY-1999 (first entry)
XX
XX Human endogenous retrovirus W clone cl.C4C5.
XX
XX Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
XX Human endogenous retrovirus.
XX
XX WO9902696-A1.
XX
XX 21-JAN-1999.
XX
XX 06-JUL-1998; 98WO-FR001442.
XX
XX 07-JUL-1997; 97FR-00008815.
XX
XX (INMR) BIO MERIEUX.
XX
XX Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX
XX WPI; 1999-120897/10.
XX
XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.
XX
XX Claim 1; Page 59-60; 106pp; French.
XX
XX PS
XX This sequence represents clone cl.C4C5 of the human endogenous retrovirus.

(HERV) W genome. The nucleic acids, their fragments or peptides encoded
by them are markers of autoimmune disease (e.g. multiple sclerosis,
rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
dependent diabetes and related pathologies) and of abnormal or
unsuccessful pregnancy and can be used as chromosomal markers for
susceptibility to these conditions, or proximity markers of genes
associated with this susceptibility

Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;

Query Match 54.3%; Score 721.2; DB 2; Length 1136;
Best Local Similarity 90.0%; Pred. No. 9.9e-226;
Matches 787; Conservative 0; Mismatches 75; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAACCCCAAAAGAGGGGAACTGTTTATTTT 60
Db 254 TCAAAATCGAAGAGCTTTAGACTTCTAACCCCAAAAGAGGGGAACTGTTTATTTT 313
QY 61 AGGGGAAGATCTGTTAGTATGTTAAATCAATCTGGAATCATCTGAGAAAGTTAAAGA 120
Db 314 AGGGGAAGATCTGTTATTTATTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 373
QY 121 AATTGAGATCGAATATATATGTTAGACGAGACCTTCAAAACACTGCACCCCTGGGCT 180
Db 374 AATTGAGATCGAATATCAACGTTAGACGAGAGCTTCGAAACACTGGACCCCTGGGCT 433
QY 181 CCTGAGCAATCGATGCGCTGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATAT 240
Db 434 CTTGAGCAATCGATGCGCTGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATAT 493
QY 241 TTTACTCTCTTTTGGACCCCTGATCTTCAACTTCTTGTAAAGTTGCTCTTCCAGAT 300
Db 494 GCTACTCTCTTTTGGACCCCTGATCTTAACTCTTGTAACTTGTCTCTTCCAGAT 553
QY 301 TGAAGCTGAAAGCTACAAATAGTTCTTCAATGNAACCCAGATCAGTCCATGACTAA 360
Db 554 CGAAGCTGTAAGCTA-----CAAAATGGAGCCCAAGATCAGTCCAAAGACTAA 601
QY 361 AATCTACCGTGGACCCCTCGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTA 420
Db 602 GATCTACCGCAGACCCCTCGACCGGCTGCTAGCCCAAGATCTGATGTTAATGACATCA 661
QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTACATCTCAATTCAGTAGG 480
Db 662 AGGCACCCCTCTCGAGGAAATCTCAGCTGCACAAACCTCTACTACGCCCCCAATTCAGCAGG 721
QY 481 AAGCAGTTAGAGAGTTGTTCAGGCAACCTCCCAACAGTACTTGGCTTTCTGTTGAGA 540
Db 722 AAGCAGTTAGAGAGTTGTTCAGGCAACCTCCCAACAGTACTTGGCTTTCTGTTGAGA 781
QY 541 GGTGGACTCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT 600
Db 782 TGGGGACTCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCT 841
QY 601 ANCTGGGAAGGTGACCGCATCTTAAACATGCGGGCTTGCACCTAGCTCAGACCCG 660
Db 842 AGCTGGGAAGGTGACCGCATCTTAAACATGCGGGCTTGCACCTAGCTCAGACCCG 901
QY 661 ACCAATCAGAGGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 902 ACCAATCAGAGGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 961
QY 721 AATCATCTATTGCTCAGAGCAGACGCGGAAGGATTTGGGATATAAACTCAGGCA 780
Db 962 AATCATCTATTGCTCAGAGCAGACGCGGAGGAGCAATGATCGGATATAAAACCAAGTC 1021
QY 781 TTCAAGCCAGCAACAGAAACCCCTTTGGGTCCCTCCCAATGATGGAGCTCTGTTTT 840
Db 1022 TTCAAGCCAGCAACAGAAACCCCTTTGGGTCCCTCCCAATGATGGAGCTCTGTTTT 1081
QY 841 CACTCTATTTCACCTATTAAATCATGCAACTGC 874
Db 1082 CATGCTATTTCACCTATTAAATCTTGCAGCTGC 1115

Db	494	GCTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCTCTTGTAACTTTGTCTCTTCCAGAAAT	553
Qy	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCAGTCCATGACTAA	360
Db	554	CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATGCAGTCCCAAGACTAA	601
Qy	361	AATCTACCGTGGACCCCTGGACCGGCTCTGTAGACTATGCTCTGTATGTTTAATGACATTGA	420
Db	602	GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTTAATGACATCA	661
Qy	421	AGTCACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG	480
Db	662	AGGCACCCCTCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG	721
Qy	481	AAGCAGTTAGACAGATTGTCAGCCAACTCCCAACAGTACTTTGGGTTTCTCTGTTGAGA	540
Db	722	AAGCAGTTAGAGCGTCTGTCGGCCAACTCCCAACAGCACTTAGGTTTCTCTGTTGAGA	781
Qy	541	GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCNAAAGCCT	600
Db	782	TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGCCT	841
Qy	601	ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660
Db	842	AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG	901
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGCGAAAAACAGAGGTAAGCAATAGCC	720
Db	902	ACCAATCAGAGAGCTCACTAAATGCTAATTAGCGAAAGACAGAGGTAAGCAATAGCC	961
Qy	721	AATCATCTATTGCTGAGACACAGCGGGAAGGACCAAGGATTTGGGATATATAACTCAGGCA	780
Db	962	AATCATCTATTGCTGAGACACAGCGGGAAGGACCAATGATCGGGATATATAACCCAAAGTC	1021
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGCTCCCTCCCATTTGATGGGAGCTCTGTTTT	840
Db	1022	TTGAGCGCGCAACGGCAACCCCTTTGGGTGCTCCCTCTCTTTGATGGGAGCTCTGTTTT	1081
Qy	841	CACCTATTTCACCTCTATTAAATCATGCAACTGC	874
Db	1082	CATGCTATTTCACCTCTATTAAATCATTCGACGTGC	1115
RESULT 25			
AAAX25661			
ID	AAAX25661 standard; cDNA to mRNA; 2782 BP.		
XX	AAAX25661;		
AC	AAAX25661;		
XX	21-MAY-1999 (first entry)		
DT	Human endogenous retrovirus W clone cl.PH74.		
DE	Clone; human endogenous retrovirus; genome; autoimmune disease;		
XX	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;		
KW	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.		
XX	Human endogenous retrovirus.		
OS	WO9902696-A1.		
PN	21-JAN-1999.		
PD	06-JUL-1998; 98WO-FR001442.		
PF	07-JUL-1997; 97FR-00008815.		
PR	(INMR) BIO MERIEUX.		
XX	Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;		
XX	WPI; 1999-120897/10.		
XX			

RESULT 24

AAAS9210

ID AAAS9210 standard; DNA; 1136 BP.

XX

AC AAAS9210;

XX

DT 07-NOV-2000 (first entry)

XX

DE 3' pol gene and 3' non coding sequences of HERV-W from human genome.

XX

KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;

KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX

OS Homo sapiens.

XX

PN WO200043521-A2.

XX

PD 27-JUL-2000.

XX

PF 21-JAN-2000; 2000WO-FR000144.

XX

PR 21-JAN-1999; 99FR-00000888.

XX

PA (INMR) BIO MERIEUX.

XX

PI Paranhos-Baccala G, Mallet F, Voisset C;

XX

DR WPI; 2000-499229/44.

XX

PT New nucleic acid from human endogenous retrovirus, useful e.g. for

PT diagnosis of autoimmune disease and complications of pregnancy, contains

PT at least part of the gag gene.

XX

PS Disclosure; Page 46; 53pp; French.

XX

CC The present sequence represents an endogenous retroviral nucleic acid

CC fragment, which is associated with an autoimmune disease, and is

CC integrated into the human genome. The fragment is originally derived from

CC a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W

CC retrovirus is associated with autoimmune disease, failure of pregnancy or

CC disorders of pregnancy. The nucleic acid fragment, or proteins derived

CC from it, are useful for diagnosis of autoimmune disease (specifically

CC multiple sclerosis) and for monitoring pregnancy. The nucleic acid

CC fragments may also be used for in situ labelling of isolated chromosomes,

CC while the transcription product can be used to study or monitor T cell

CC proliferation in vitro

XX

SQ Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;

Qy	1	TCAAATCGAAGCTTTAGACTTGCTTAACCGCAAGAGGGGAACTGTTATTTT	60
Db	254	TCAAATCGAAGCTTTAGACTCGCTAACCGCTGAAGAGGGGAACTGTTATTTT	313
Qy	61	AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAATTAAGA	120
Db	314	AGGGGAAGATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAATTAAGA	373
Qy	121	AATTTGAGATCGAATATAATGTAGACGAGGACCTTCAAAACACTGCACCTGGGGCT	180
Db	374	AATTCGAGATCGAATACACGTAGACGAGGAGCTTCGAAACACTGGACCTGGGGCT	433
Qy	181	CCTCAGCAATGATGCTCGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATATATT	240
Db	434	CCTCAGCAATGATGCTCGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATATATT	493
Qy	241	TTTACTCCTCTTTGGACCCCTGTATCTTCAACTCTCTTTGTTAGTTTGTCTCTCCAGAT	300

PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
 PT disease, and abnormal or failed pregnancy.

XX Claim 1; Page 60-63; 106pp; French.

XX This sequence represents clone cl.PH74 of the human endogenous retrovirus
 CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
 CC rheumatoid polyarthritits, disseminated lupus erythematosus, insulin-
 CC dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility

XX Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 53.6%; Score 712.6; DB 2; Length 2782;

Best Local Similarity 89.4%; Pred. No. 1.1e-222;

Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTGTCTAAACCGCTGAGAGAGGGGAACTGTTATTTT 60
 DB 1902 TCGAAATCGAAGAGCTTTAGACTGTCTAAACCGCTGAGAGAGGGGAACTGTTATTTT 1961

QY 61 AGGGGAAGATCTGTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
 DB 1962 AGGGGAAGATCTGTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGA 2021

QY 121 AATTTCAGATCGAATATAATGTAGACAGAGGACCTTCAAAACACTGCACCTCGGGCCCT 180
 DB 2022 AATTTCAGATCGAATACACGATATAGACAGAGGAGCTTCGAAACACTGGACCTCGGGCCCT 2081

QY 181 CCTCAGCCAAATGATGCGCTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT 240
 DB 2082 CCTCAGCCGATGATGCGCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT 2141

QY 241 TTTTACTCTCTTTGGACCTGTATCTCAATCTTCTTGTAGTTTGTCTCTTCCAGAAAT 300
 DB 2142 GCTACTCTCTTTGGACCTGTATCTTGGACCTCTTGTAACTTTGTCTCTTCCAGAAAT 2201

QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAATGGAAACCCAGATGAGTCCATGACTAA 360
 DB 2202 CGAAGCTGTAAACTA-----CAAATGGAGCCCAAGATGAGTCCAAAGACTAA 2249

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTATGTTTAATGACATTTGA 420
 DB 2250 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCCAACGATCTGATGTTAATGACATCAA 2309

QY 421 AGTCACCCCTCCCGAGGAATCTCACTGCACACCCCTACTACCTCCCAATTCAGTAGG 480
 DB 2310 AGGCACCCCTCCCTGAGGAATCTCAGCTGCACACCTCTACTACGCCCAATTCAGCAGG 2369

QY 481 AAGCAGTTAGACAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCCTGTTGAGA 540
 DB 2370 AAGCAGTTAGAGGTTGTTCGGCCAACTCCCAACAGCATTAGTTTCTGTTGAGA 2429

QY 541 GGTGTGACTGAGAGACAGGACTAGCTGGAATTTCTTAGGCTGACTAAGATTCNNAGCCT 600
 DB 2430 TGGGGGACTGAGAGACAGGACTAGCTGGAATTTCTTAGGCTGACTAAGATTCNTTAAGCCT 2489

QY 601 ANCTGGAGGTCACCGCATCTTAAACATGGGCTTGAACCTTAGCTCAGACCCG 660
 DB 2490 AGTGGGAAGGTGACCAATCCACCTTTAAACACGCGGCTTGAACCTTAGCTCAGACCTG 2549

QY 661 ACCAATCAGAGAGCTCACTAAATATGCTAATCAGGCAAAACACAGAGGTAAGCAATAGCC 720
 DB 2550 ACCAATCAGAGAGCTCACTAAATATGCTAATAGGCAAGACAGAGGTAAGCAATAGCC 2609

QY 721 AATCATCTATTGCTGAGACACAGCGGGAAGGACAAGAGATTGGGATATATAAATCAGGCA 780
 DB 2610 AATCATTTATTGCTGAGACACAGCAGGAGGAGCAATGATCGGGATATATAAATCAGGAT 2669

QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTTT 840
 DB 2670 TTGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTTT 2729

QY 841 CACTCTATTCTACTCTATTAAATCATGCAACTGCA 875

DB 2730 CATGCTATTCTACTCTATTAAATCTTGCAACTGCA 2764

RESULT 26

AAA59211

ID AAA59211 standard; DNA; 2782 BP.

AC AAA59211;

XX 07-NOV-2000 (first entry)

XX 5' non coding, 3' pol gene, and 3' non coding sequences of HERV-W.

XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;

XX gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX Homo sapiens.

XX WO200043521-A2.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-FR000144.

XX 21-JAN-1999; 99FR-00000888.

XX (INMR) BIO MERIEUX.

XX Paranhos-Baccala G, Mallet P, Voisset C;

XX WPI; 2000-499229/44.

XX New nucleic acid from human endogenous retrovirus, useful e.g. for

XX diagnosis of autoimmune disease and complications of pregnancy, contains

XX at least part of the gag gene.

XX Disclosure; Page 46-47; 53pp; French.

XX The present sequence represents an endogenetic retroviral nucleic acid
 fragment, which is associated with an autoimmune disease, and is
 integrated into the human genome. The fragment is originally derived from
 a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
 retrovirus is associated with autoimmune disease, failure of pregnancy or
 disorders of pregnancy. The nucleic acid fragment, or proteins derived
 from it, are useful for diagnosis of autoimmune disease (specifically
 multiple sclerosis) and for monitoring pregnancy. The nucleic acid
 fragments may also be used for in situ labelling of isolated chromosomes,
 while the transcription product can be used to study or monitor T cell
 proliferation in vitro

XX Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 53.6%; Score 712.6; DB 3; Length 2782;

Best Local Similarity 89.4%; Pred. No. 1.1e-222;

Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTGTCTAAACCGCAAAAGAGGGGAACTGTTATTTT 60
 DB 1902 TCGAAATCGAAGAGCTTTAGACTGTCTAAACCGCTGAGAGGGGAACTGTTATTTT 1961

QY 61 AGGGGAAGATCTGTTAGTATGTTAATCAATCTGGAATCATTTCTGAGAAAGTTAAAGA 120
 DB 1962 AGGGGAAGATCTGTTATTTATCTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGA 2021

QY 121 AATTTCAGATCGAATATAATGTAGACAGAGGACCTTCAAAACACTGCACCTCGGGCCCT 180
 DB 2022 AATTTCAGATCGAATACACGATATAGCAGAGGAGCTTCGAAACACTGCACCTCGGGCCCT 2081

QY 181 CCTCAGCAATGGATGCCCTGGACTCCCTCTTAGGACCTCTAGCAGCTAATAATT 240
Db 2082 CCTCAGCGATGGATGCCCTGGACTCCCTCTTAGGACCTCTAGCAGCTAATAATT 2141
QY 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTTCTTGGTTAAAGTTTGTCTCTCCAGAAAT 300
Db 2142 GCTACTCTCTTTGGACCCCTGTATCTTGGACTCTCTTGTAACTTTGTCTCTCCAGAAAT 2201
QY 301 TGAAGCTGTAAGCTTACAATAAGTTCTTCAATAGGAACCCAGATGAGTCCATGACTAA 360
Db 2202 CGAAGCTGTGAACATA-----CAAATGGAGCCCAAGATGAGTCCCAAGACTAA 2249
QY 361 AATCTACCGTGGACCCCTGGACCCGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db 2250 GATCTACCGCAGACCCCTGGACCCGCTGTAGCCACGATCTGATGTTAATGACATCA 2309
QY 421 AGTCACCCCTCCGAGGAATCTCACTGCAACACCCCTACTACACTCCAATTCAGTAGG 480
Db 2310 AGGCACCCCTCCTGAGGAATCTCAGCTGCAACACCTCTACTAGCCCAATTCAGCAGG 2369
QY 481 AAGCAGTTTAGACAGCTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTGAGA 540
Db 2370 AAGCAGTTTAGACAGCTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTGAGA 2429
QY 541 GGGTGGACTGAGACAGACGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCNAAGCCT 600
Db 2430 TGGGGGACTGAGACAGACGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCTTAAGCCT 2489
QY 601 ANCTGGGAGGTGACCGCATCATCTTTAAACATGGGCTTGCACTTAGCTCACACCG 660
Db 2490 AGGTGGGAGGTGACCATCATCTTTAAACATGGGCTTGCACTTAGCTCACACCTG 2549
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2550 ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCC 2609
QY 721 AATCATCTATTCCTGAGACGACGAGGAGCAAGATTTGGATATTAATCACTCAGCA 780
Db 2610 AATCATTTATTCCTGAGACGACGAGGAGCAAGATTTGGATATTAATCACTCAGCA 2669
QY 781 TTCAGGCGGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTGTT 840
Db 2670 TTCGAGCGGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTGTT 2729
QY 841 CACTCTATTCTACTTATTAAATCATGCAACTGCA 875
Db 2730 CATGCTATTCTACTTATTAAATCTTGCAACTGCA 2764

RESULT 27
AAH20069
ID AAH20069 standard; DNA; 2782 BP.
XX AAH20069;
AC
XX
DT 08-AUG-2001 (first entry)
XX
DE HERV-W envelope protein G encoding nucleic acid.
XX
KW Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW envelope protein; multiple sclerosis-related superantigen; vaccine;
KW surface antigen; transmembrane; multiple sclerosis; neuroprotective;
KW antisense-therapy; autoimmune disorder; ds.
XX
OS Human endogenous retrovirus.
XX
FH Key Location/Qualifiers
FT 5' UTR 1..762 a
FT 763..2379
FT CDS /*tag= b
FT /*product= "HERV-W envelope protein G"
FT

FT /transl_except= (pos:790..792,aa:Phe)
FT /transl_except= (pos:793..795,aa:Thr)
FT /transl_except= (pos:812..814,aa:Leu)
FT /transl_except= (pos:818..820,aa:Ser)
FT /transl_except= (pos:862..864,aa:Tyr)
FT /transl_except= (pos:865..867,aa:Gln)
FT /transl_except= (pos:1174..1176,aa:Arg)
FT /transl_except= (pos:1441..1443,aa:Leu)
FT /transl_except= (pos:1903..1905,aa:Gln)
FT /transl_except= (pos:2017..2019,aa:Lys)
FT /transl_except= (pos:2026..2028,aa:Arg)
FT /transl_except= (pos:2044..2046,aa:Arg)
FT /transl_except= (pos:2089..2091,aa:Gln)
FT /transl_except= (pos:2170..2172,aa:Asn)
FT 2380..2782
FT /*tag= c
XX
XX W0200131021-A1.
XX
XX 03-MAY-2001.
XX
XX 30-OCT-2000; 2000WO-EP010659.
XX
XX 28-OCT-1999; 99EP-00402690.
XX (UYGE-) UNIV GENEVE.
XX
XX Conrad B, Mach B;
XX
XX WPI; 2001-316336/33.
XX P-PSDB; AAB75138.
XX
XX New human retrovirus HERV-W ENV proteins/peptides having superantigen
XX activity useful for diagnosing and treating multiple sclerosis.
XX
XX Claim 13; Fig 9; 94pp; English.
XX
XX On the basis of the PBS t-RNA motif used for the classification of human
XX endogenous retrovirus (HERVs) the full length endogenous provirus which
XX was located on the long arm of human chromosome 7 (7q21-22) has been
XX designated HERV-W. The present invention describes proteins or peptides
XX (I) having superantigen (SAG) activity comprising the ENV protein (ENV)
XX of HERV-W, the surface protein (SU) and can be used in: vaccines; antisense-
XX therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are
XX useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
XX disorders. (I) are also useful for identifying superantigen associated
XX recovering) capable of binding to a retroviral superantigen and optionally
XX with MS, substances capable of blocking SAG activity and substances
XX capable of blocking transcription or translation of HERV-W retroviral
XX superantigen. A protein or peptide derived from (I), modified to be
XX devoid of SAG activity and being capable of generating an immune response
XX against HERV-W retroviral SAG is useful in therapy. Nucleic acid
XX molecules encoding (I) are useful as vaccines against MS. Substances
XX capable of blocking SAG activity, capable of binding to a retroviral
XX superantigen associated with MS, or capable of blocking transcription or
XX translation of HERV-W retroviral superantigen for use in treating or
XX preventing MS, obtained using (I) are useful for the treatment and
XX prevention of MS. (I) and nucleic acids encoding them are useful for
XX diagnosing autoimmune disease. The present sequence encodes the
XX specifically claimed envelope protein of HERV-W designated G
SQ Sequence 2782 BP; 741 A; 768 C; 564 G; 709 T; 0 U; 0 Other;
Query Match 53.6%; Score 712.6; DB 5; Length 2782;
Best Local Similarity 89.4%; Pred. No. 1.1e-222;
Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAACCAGCCAAAGAGGGGAACTGTTATTTT 60
Db 1902 TCGAAATCGAAGAGCTTTAGACTTCTAACCAGCCGCTGAGAGGGGAACTGTTATTTT 1961
QY 61 AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCACTACTGAGAAAGTTAAAGA 120

Db 1962 AGGGAAGAAATCTGTTATATATGTTAATCAATCCGGAATCGTCACTGAGAAATTTGAAGA 2021
Qy 121 AATTTGAGATCGAATATATATAGAGCAGAGGACCTTCAAAACACTGCAACCTGGGGCT 180
Db 2022 AATTCAGATCGAATACAAACGATATAGCAGAGGAGCTTCAAAACACTGCAACCTGGGGCT 2081
Qy 181 CCTCAGCAATGATGCGCTGCACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 2082 CCTCAGCCGATGATGCGCTGCACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 2141
Qy 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTAAAGTTTGTCTCTTCCAGAT 300
Db 2142 GCTACTCTCTTTGGACCTGATCTTGTGACCTCTTGTAACTTGTCTCTTCCAGAT 2201
Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAAACCCAGATGCAATGCACTAA 360
Db 2202 CGAAGCTGTAACTA-----CAAATGGAGCCCAAGATGCAATGCACTAA 2249
Qy 361 AATCTACCTGAGACCTCGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2250 GATCTACCGCAGACCTCGACCGGCTGCTAGCCACCATCTGATGTTAATGACATCAA 2309
Qy 421 AGTCACCTCCCGAGGAAATCTCAACTGCACAAACCTTACTACACTTCAATTCAGTAGG 480
Db 2310 AGGCACTCTCTGAGGAAATCTCAGCTGCACAACTTACTAGCCCAATTCAGCAGG 2369
Qy 481 AAGCAGTTAGCAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
Db 2370 AAGCAGTTAGCGGTGGTGGCCAACTCCCAACAGCAGCTTAGGTTTCTGTTGAGA 2429
Qy 541 GGGTGAAGTGAAGCAGACTAGCTGGATTCTTAGGCTGACTAAGATCCCNAGCCT 600
Db 2430 TGGGGGACTGAGAGCAGGACTAGCTGGATTCTTAGGCTGACTAAGATCCCTTAGCCT 2489
Qy 601 ANCTGGGAAGTGCACCGCATCTTAAACATGGGGCTTGCACCTTAGCTCACACCCG 660
Db 2490 AGTGGGAAGTGCACATCTCACTTAAACAGGGGCTTGCACCTTAGCTCACACCTG 2549
Qy 661 ACCAATCAGAGAGCTCACTAAAATGTCTAATCAGGCAAAACAGAGGTTAAGCAATAGCC 720
Db 2550 ACCAATCAGAGAGCTCACTAAAATGTCTAATCAGGCAAAACAGAGGTTAAGCAATAGCC 2609
Qy 721 AATCATCTATTGCTGAGACACAGCGGAGGACAGGATTTGGGATATAAATCAGGCA 780
Db 2610 AATCATTTATTGCTGAGACACAGCGGAGGACAAATGATCGGGATATAAATCAGGAT 2669
Qy 781 TTCAAGCCAGCAACAGCAACCTTTGGGTCCCTCCCAATTTGATGGGAGCTCTGTTTT 840
Db 2670 TTGAGCGGCAACGGCAACCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTT 2729
Qy 841 CACTCTATTTTCACTTATTAATCATGCAACTGCA 875
Db 2730 CATGCTATTTCACTTATTAATCTTGCAACTGCA 2764

RESULT 28

AA114608

ID AA114608 standard; DNA; 1894 BP.

XX AC AA114608;

XX AC AA114608;

XX AC AA114608;

XX AC AA114608;

XX AC AA114608;

DE DE Probe #4541 for gene expression analysis in human cervical cell sample.

XX XX Probe; human; microarray; gene expression; cervical epithelial cell;

XX XX cervical cancer; ss.

XX XX Homo sapiens.

XX XX Homo sapiens.

XX XX Homo sapiens.

XX XX Homo sapiens.

XX XX Homo sapiens.

XX XX Homo sapiens.

PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000670.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX Claim 25; SEQ ID NO 4541; 487bp; English.
XX The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human Hela cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
SQ Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 120
Db 961 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 1080
Qy 61 AGGGGAAGATCTGTTAGTATGTTAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 120
Db 1021 AGAGGAAATGCTGTTGTTAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 1080
Qy 121 AATTTGAGATCGAATATAATGTAGACGAGGACCTTCAAAACACTGCAACCTGGGGCCT 180
Db 1081 AATTTGAGATCGAATATAATGTAGACGAGGACCTTCAAAACACTGCAACCTGGGGCCT 1140
Qy 181 CCTCAGCAATGATGCGCTGCACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 1141 CCTCAGCAATGATGCGCTGCACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 1200
Qy 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTAAAGTTTGTCTCTTCCAGAT 300
Db 1201 GTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTAAAGTTTGTCTCTTCCAGAT 1260
Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAAACCCAGATGCAATGCACTAA 360
Db 1261 CGAAGCAGTAAACTACAAATCGTTCTTCAATGGAGCCCAAGATGCAATGCACTAA 1320
Qy 361 AATCTACCTGAGACCTCGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 1321 AATCTACCTGAGACCTCGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 1380
Qy 421 AGTCACCTCTCCGAGGAAATCTCAACTGCACAAACCTTACTAGCCCAATTCAGTAGG 480
Db 1381 AGGCACCTCTCCGAGGAAATCTCAACTGCACAAACCTTACTAGCCCAATTCAGTAGG 1440
Qy 481 AAGCAGTTAGCAGGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTGCTTAAACCGCAAAAGAGGGGACCTGTTTATTTT 60
Db 961 TCAAAATCGAAGAGCTTTAGACTTGTGCTTAAACCGCAAAAGAGGGGACCTGTTTATTTT 1020
Qy 61 AGGGGAGAGTGTCTGTATGTTTAAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
Db 1021 AGAGGAAAAATGCTGTGTGTTTATGTTTAAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 1080
Qy 121 AATTTGAGATCGAATATAATGTAGACGAGAGACCTTCAAAAACATGACACCTGGGCGCT 180
Db 1081 AATTCAGGTGGAATATAACGTAGAGCAAGAGAGCTGCAAAAACATGACACCTGGGCGCT 1140
Qy 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTCTTTAGGACCTCTAGCAGCTATAATTT 240
Db 1141 CCTCAGCAATGGATGCCCTGGACTCTCCCTTCTTTAGGACCTCTAGCAGCTATAATTT 1200
Qy 241 TTTTACTCTCTTTGGACCTGTATCTTCAAACTTCTTGTAAAGTTTCTTCTCCAGAAT 300
Db 1201 GTTACTCTCTTTGGACCTGTATCTTCAAACTTCTTGTAAAGTTTCTTCTCCAGAAT 1260
Qy 301 TGAAGCTGTAAGACTCAAAATAGTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
Db 1261 CGAAGCAGTAAACTCAAAATCGTTCTTCAAAATGGAACCCAGATGAGTCCATGAGTAA 1320
Qy 361 AATCTACGTTGGACCTCGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTTGA 420
Db 1321 AATCTACGTTGGACCTCGACCGGCTGTAGCCCATGCTCTGATGTTAATGACATCAA 1380
Qy 421 AGTCACCTCTCCCGAGGAAATCTCAACTGCAACACCCCTACTACCTCCAAATTCAGTAGG 480
Db 1381 AGGCACCTCTCCCGAGGAAATCTCAACTGCAACACCTCTACTACGCCCCCAATTCAGCAGG 1440
Qy 481 AAGCAGTTAGACGAGTTGTGACGCCAACCTCCCAACAGTACTTGGGTTTCTCTGTGAGA 540
Db 1441 AAGCAGTTAGAGTGTGTTGTTGGCCAACTCTCCCAACAGTACTTGGGTTTCTCTGTGAGA 1500
Qy 541 GGGTGGACTGAGACAGGAGTACTGATGATTTCTTAGGCTGACTAAGAAATCCNAAGCCT 600
Db 1501 GGGGGAATGAGACAGGAGTAACTAGATTTCTTAGACCACTAAGAAATCCCTAAGACT 1560
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCACCTTTAGCTACACCG 660
Db 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCACCTTTAGCTACACGCCA 1620
Qy 661 ACCAATC-----AGAGAGCTCACTAAATGCTATCAGGCAAAACAGGAGGTAAA 711
Db 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTATTAAGGCAAAACAGGAGATAAA 1680
Qy 712 GCAATAGCCATCATCTATTGCTGAGACGACAGCGGGAAGGACAAGATTGGGATATAA 771
Db 1681 GAAATAGCCATCATCTGTTGCTTGACAGACAGCAGGAGGAGCAATGATCGGATATAA 1740
Qy 772 ACTCAGGCATTTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGGAG 831
Db 1741 ACCCAGGCATTTGAGCCAGCTACAGCTCTTTGGGTCCCTCCCTTCTTGTATGGGAG 1800
Qy 832 CTCGTCTTCACTCTATT 849
Db 1801 CTCGTCTTCACTCTATT 1818
RESULT 29
ID ABA56337 standard; DNA; 1894 BP.
XX ABA56337;
XX ABA56337;
DT 01-FEB-2002 (first entry)
XX Human foetal liver single exon nucleic acid probe #4642.
XX Human foetal liver; gene expression; single exon nucleic acid probe; ss.
XX Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human foetal liver.
XX Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed

Db 1441 AAGCAGTTAGAGTGTGTTGTTGCGCAACCTCCCAACAGCAGTGGGTTTCTCTTTGAGA 1500
Qy 541 GGGTGGACTGAGACGAGTACTGATGTTTCTTAGGCTGACTAAGAAATCCNAAGCCT 600
Db 1501 GGGGGGACTGAGAGACAGGAATAACTAGATTTCTTAGACCAACTAAGAAATCCCTAAGACT 1560
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCACCTTTAGCTACACCG 660
Db 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCACCTTTAGCTACACGCCA 1620
Qy 661 ACCAATC-----AGAGAGCTCACTAAATGCTATCAGGCAAAACAGGAGGTAAA 711
Db 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTATTAAGGCAAAACAGGAGATAAA 1680
Qy 712 GCAATAGCCATCATCTATTGCTGAGACGACAGCGGGAAGGACAAGATTGGGATATAA 771
Db 1681 GAAATAGCCATCATCTGTTGCTTGACAGACAGCAGGAGGACANATGATCGGATATAA 1740
Qy 772 ACTCAGGCATTTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGGAG 831
Db 1741 ACCCAGGCATTTGAGCCAGCTACAGCTCTTTGGGTCCCTCCCTTCTTGTATGGGAG 1800
Qy 832 CTCGTCTTCACTCTATT 849
Db 1801 CTCGTCTTCACTCTATT 1818
RESULT 29
ID ABA56337 standard; DNA; 1894 BP.
XX ABA56337;
XX ABA56337;
DT 01-FEB-2002 (first entry)
XX Human foetal liver single exon nucleic acid probe #4642.
XX Human foetal liver; gene expression; single exon nucleic acid probe; ss.
XX Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human foetal liver.
XX Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed

XX AAI35980;
 AC 17-OCT-2001 (first entry)
 XX Probe #4666 used to measure gene expression in human placenta sample.
 XX DE Probe #4666 used to measure gene expression in human placenta sample.
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder; ss.
 XX OS Homo sapiens.
 XX WO200157272-A2.
 XX PD 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000663.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX PI WPI; 2001-48897/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human placenta.
 XX Claim 25; SEQ ID NO 4666; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP).
 XX The present sequence is one such probe. The probes are useful for
 XX producing a microarray for predicting, measuring and displaying gene
 XX expression in samples derived from human placenta. The probes are useful
 XX for antenatal diagnosis of human genetic disorders
 XX SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 Query Match 53.0%; Score 704.8; DB 4; Length 1894;
 Best Local Similarity 89.7%; Pred. No. 3.3e-220;
 Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATTACTGAGAAAGCTTAAAGA 120
 DB 961 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATTACTGAGAAAGCTTAAAGA 1080
 QY 61 AGGGGAAGATGCTGTAGTATCTTAATCAATCTGGAATCATTACTGAGAAAGCTTAAAGA 120
 DB 1021 AGAGGAAATGCTGTGTATGTTAATCAATCTGGAATCATTACTGAGAAAGCTTAAAGA 1080
 QY 121 AATTGAGATCGAATATAATGTAGACGAGGACCTTCAAAACACTGCAACCTGGGCGCT 180
 DB 1081 AATTCAAGTCGAATATAACGTAGACAAAGGAGCTGCAAAACACTGCAACCTGGGCGCT 1140
 QY 181 CCTCAGCAATGATGCCCTGGAGCTTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 240
 DB 1141 CCTCAGCAATGATGCCCTGGAGCTTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 1200
 QY 241 TTTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAGTTGTCTCTTCCAGAAT 300
 DB 1201 GTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAGTTGTCTCTTCCAGAAT 1260
 QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
 DB 1261 CGAAGCAGTAAACTACAAATCGTCTTCAAAATGGAACCCAGATGAGTCCATGAGTAA 1320

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGGA 420
 DB 1321 AATCTACCGACCGACCCCTGGACCGGCTGTAGCCATGCTCTGATGTTAATGACATCAA 1380
 QY 421 AGTCACCCCTCCGAGGAAATCTCACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
 DB 1381 AGGCACCCCTCCGAGGAAATCTCACTGCAACCCCTACTACACTCCAATTCAGTAGG 1440
 QY 481 AAGCAGTTAGACGAGTGTCTGAGCCAACTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
 DB 1441 AAGCAGTTAGAGTGTGTGTCGCACTCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 1500
 QY 541 GGGTGGACTGAGAGACGAGTGTGTTCTTAGCTGATGTTCTTAGCTGATGATCCNAAGCCT 600
 DB 1501 GGGGGGACTGAGAGACGAGTGTGTTCTTAGCTGATGTTCTTAGCTGATGATCCNAAGCCT 1560
 QY 601 ANCTGGGAAGTGAACGATCCATCTTTAAACATGGGCTTGCACACTTACACTGCACACCCG 660
 DB 1561 AGCTGGGAAGTGAACGATCCATCTTTAAACATGGGCTTGCACACTTACACTGCACACCCG 1620
 QY 661 ACCAATC-----AGAGAGCTCACTAAATATGCTAATCAGGCAAAAACAGGAGTAAA 711
 DB 1621 ACCAATCAGATCTAAAGAGAGCTCACTAATATGCTAATTAGGCAAAAACAGGAGTAAA 1680
 QY 712 GCAATAGCAATCATCTATTGCTGAGAGCACAGCGGGAAGCAAGGATTCGGATATAA 771
 DB 1681 GAAATAGCAATCATCTATTGCTGAGAGCACAGCGGGAAGCAAGGATTCGGATATAA 1740
 QY 772 ACTCAGGCAATCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCTCCCTTATGGGAG 831
 DB 1741 ACCCAGGCAATTCGAGCCAGCTACAGTACCTCTTTGGTCCCTCCCTTATGGGAG 1800
 QY 832 CTCTGTTTCACTCTATT 849
 DB 1801 CTCTGTTTCACTCTATT 1818
 RESULT 31
 ABM45822
 ID ABM45822 standard; DNA; 1894 BP.
 XX AC ABA45822;
 XX AC ABA45822;
 XX 01-FEB-2002 (first entry)
 XX Human breast cell single exon nucleic acid probe #4517.
 XX DE Human; microarray; single exon probe; gene expression; breast; disease;
 XX KW cancer; ss.
 XX OS Homo sapiens.
 XX WO200157271-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000662.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX PI WPI; 2001-496933/54.
 XX New spatially-addressable set of single exon nucleic acid probes, useful

PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 1; SEQ ID NO 4517; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. NO. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;

QY 1 TCMAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACCTGTTATTTT 60
DB TCMAAATCGAAGAGCTTTAGACTTCTAAACCGCGGAAGCGGGGAACCTTTTATTTT 1020

QY 61 AGGGAAGAAGTCTGTAGTATGTTAATCAATCGGAATCAATCTGAGAAGTTAAAGA 120
DB AGAGGAAATAGCTGTGTGTATGTTAATCAATCGGAATCAATCCGGAAGTTAAAGA 1080

QY 121 AATTGAGATCGAATATAATGATAGACAGAGGACCTTCAAAACACTGCACCTGGGCT 180
DB AATTCAAGGTCGAATATAACGTAGACAGCAAGAGCTGCANAACACTGGACCTGGGCT 1140

QY 181 CCTCAGCAATGAGTCCCTGACCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
DB CCTCAGCAATGAGTCCCTGAGTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 1200

QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTGTTAAGTTTGTCTTTCAGAAAT 300
DB GTTACTCTCTTTGGACCTGATCTTCAACTTCTGTTAAGTTTGTCTTTCAGAAAT 1260

QY 301 TGAAGCTGTAAGCTTACAATAGTTCTTCAATGGAACCCAGATGCAATGACTAA 360
DB CGAAGCAGTAAACTACAAATCGTTCTTCAATGGAAGCCAGATGCAATGACTAA 1320

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTGA 420
DB AATCTACCGGACCCCTGGACCGGCTGCTAGCCATGCTCTGATGTTAATGACTAA 1380

QY 421 AGTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTACACTCAATTCAGTAGG 480
DB AGGCACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTACCGCCCAATTCAGCAGG 1440

QY 481 AAGCAGTTAGACAGTCTGACCAACCTCCCAACAGTCTTGGGTTTCTGTTGAGA 540
DB AAGCAGTTAGAGTGTGTTGGCCAACTCCCAACAGAGTGGGTTTCTGTTGAGA 1500

QY 541 GGGTGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCT 600
DB GGGGGGACTGAGACAGGAAATACTAGATTTCTTAGCAACCACTAAGAAATCCCTAAGACT 1560

QY 601 ANCTGGGAAGGTGACCGCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCG 660
DB AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACGCCA 1620

QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGTAAA 711
DB 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTAATTAGCAAAACAGGAGATAA 1680

QY 712 GCAATAGCCCAATCATCTATTTCCTGAGAGCACAGCGGGAAGGACAGGATTGGGATATA 771
DB 1681 GAAATAGCCCAATCATCTATTTCCTGAGAGCACAGGAGGGAACAATATCGGGATATA 1740

QY 772 ACTCAGCATTCAAGCCCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTTATGGAG 831
DB 1741 ACCAGGCATTCGAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGTTATGGAG 1800

QY 832 CTCTGTTTTCACTCTATT 849
DB 1801 CTCTGTTTTCACTCTATT 1818

RESULT 32
ABA25978
ID ABA25978 standard; DNA; 1894 BP.
XX
AC ABA25978;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #4444 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
OS Homo sapiens.
XX
FN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human hearts.
XX
PS Claim 1; SEQ ID NO 4444; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTTATTTT 60
Db 961 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCGAAAGCGGGGAACTTTTATTTT 1020

QY 61 AGGGGAAGATGCTGTAGTATGTTAATCAATCTGAATCATTTACTGAGAAAGTTAAAGA 120
Db 1021 AGAGGAAAATGCTGTGTATGTTAATCAATCCGGAATCATCACCAGAAAAGTTAAAGA 1080

QY 121 AATTTGAGATCGAATATAATGTAGAGCAGAGGACCTTCAAAACACTGCACCCCTGGGGCCT 180
Db 1081 AATTCAGGTGCAATATACGTAGAGCAAGAGCTGCAAAACACTTGGACCCCTGGGGCCT 1140

QY 181 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 1141 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

QY 241 TTTTACTCCTTTTGGACCCCTGTATCTTCACTTCTGTTTAAAGTTTGTCTTCCAGAA 300
Db 1201 GTTACTCCTTTTGGACCCCTGTATCTTAACTTCTGTTTAAAGTTTGTCTTCCAGAA 1260

QY 301 TGAAGCTGTAAGCTCAAAATAGTCTTCAAAATGGAACCCAGATCAGTCCATGACTAA 360
Db 1261 CGAAGCAGTAAACTCAAAATAGTCTTCAAAATGGAACCCAGATCAGTCCATGACTAA 1320

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 420
Db 1321 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 1380

QY 421 AGTCACCCCTCCGAGGAATCTCAACTGACAGAACCCCTACTACTCTCACTCCAAATTCAGT 480
Db 1381 AGGCACCCCTCCGAGGAATCTCAACTGACAGAACCCCTACTACTCTCACTCCAAATTCAGT 1440

QY 481 AAGCAGTTAGAGGTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
Db 1441 AAGCAGTTAGAGGTTGTTGGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA 1500

QY 541 GGGTGGACTGAGAGCAGGACTAGTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 600
Db 1501 GGGGGGACTGAGAGCAGGAAATTAAGTATTTCTTAGACCACTAAGAAATCCCTAAGACT 1560

QY 601 ANCTGGAGGTTGACCGCATCTTTTAAACATGGGCTTGCACCTAGCTCAGACCCG 660
Db 1561 AGCTGGGAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCACCTAGCTCAGCCCA 1620

QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTAAATCAGGCAAAAACAGGAGTAAA 711
Db 1621 ACCAATCAGATCTAAGAGAGCTCACTAAATGCTAAATAGGCAAAAACAGGAGTAAA 1680

QY 712 GCAATAGCCAATCATCTATTGCTGAGACACAGCGGGAAGGACAAGGATTTGGGATATA 771
Db 1681 GAATAGCCAATCATCTATTGCTGAGACACAGCGGGAAGGACAATGATCGGATATA 1740

QY 772 ACTCAGCATTTAAGCCAGCAACAGCAACCCCTTTGGGTTCCCTCCCTATGATGGAG 831
Db 1741 ACCCAGCATTTGAGCCAGCTACAGCTACCTCTTTGGGTTCCCTCCCTTTGATGGAG 1800

QY 832 CTCTGTTTTCACCTATT 849
Db 1801 CTCTGTTTTCACCTATT 1818

RESULT 33
AAK30018
ID AAK30018 standard; DNA; 1894 BP.
XX
AC AAK30018;
XX
DT 06-NOV-2001 (first entry)
XX

DE Human bone marrow expressed single exon probe SEQ ID NO: 4575.
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
OS Homo sapiens.
XX WO200157276-A2.
PN 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000668.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-0063236P.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 4575; 658pp + Sequence Listing; English.
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTTATTTT 60
Db 961 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGGAAAGCGGGGAACTTTTATTTT 1020

QY 61 AGGGGAAGATGCTGTAGTATGTTAATCAATCTGAATCATTTACTGAGAAAGTTAAAGA 120
Db 1021 AGAGGAAAATGCTGTGTATGTTAATCAATCCGGAATCATCACCAGAAAAGTTAAAGA 1080

QY 121 AATTTGAGATCGAATATAATGTAGAGCAGAGGACCTTCAAAACACTGCACCCCTGGGGCCT 180
Db 1081 AATTCAGGTGCAATATAACGTAGAGCAAGAGCTGCAAAACACTTGGACCCCTGGGGCCT 1140

QY 181 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 1141 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

QY 241 TTTTACTCCTTTTGGACCCCTGTATCTTCACTTCTGTTTAAAGTTTGTCTTCCAGAA 300
Db 1201 GTTACTCCTTTTGGACCCCTGTATCTTAACTTCTGTTTAAAGTTTGTCTTCCAGAA 1260

QY 301 TGAAGCTGTAAGCTCAAAATAGTCTTCAAAATGGAACCCAGATCAGTCCATGACTAA 360
Db 1261 CGAAGCAGTAAACTCAAAATAGTCTTCAAAATGGAACCCAGATCAGTCCATGACTAA 1320

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 420
Db 1321 AATCTACCGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 1380

Qy	421	AGTCA	CCCCCTCCCGAGGAATCTCAACTGCAGAACCCCTACTACTACATCCCAATTCAGTAGG	480
Db	1381	AGGCAC	CCCCCTCCCGAGGAATCTCAACTGCAGAACCCCTACTACTAGCGCCCCCAATTCAGCAGG	1440
Qy	481	AAGCAG	TTAGAGCAGTTGTTCAGCGCAACCTCCCCCAAACAGTACTTTGGGTTCCTCTGTTGAGA	540
Db	1441	AAGCAG	TTAGAGTGGTTGTTGGCCAACTCCCCCAAACAGCAGTTGGGTTTCTCTGTTGAGA	1500
Qy	541	GGGTGGA	CTCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCNAAAGCCT	600
Db	1501	GGGGGGA	CTCAGAGACAGGAAATAACTAGATTTCTTAGACCAACTAAGAATCCCTAAGACT	1560
Qy	601	ANCTGGG	AAGTGCACCGCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCG	660
Db	1561	AGCTGGG	AAGTGCACCGCTTCCACCTTTAAACACCGGCTTGCAACTTAGCTCACGCCCA	1620
Qy	661	ACCAATC	-----AGAGAGCTCACTAAAATGCTAATCAGGCAAAAAACAGGAGGTAAA	711
Db	1621	ACCAATC	CAGATACTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAAACAGGAGATAA	1680
Qy	712	GCAATAG	CCCAATCTATTGCTTGAGAGCACAGCGGGAAGGACAAGGATTCGGATATAA	771
Db	1681	GAATAGC	CAATCATCTGTTGCCCTGCAGCACAGCAGGAGGGAACAATGATCGGGATATAA	1740
Qy	772	ACTCAGG	CATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAG	831
Db	1741	ACCAGCA	CTTCAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGATGGAG	1800
Qy	832	CTCTG	TTTTCACTCTATT	849
Db	1801	CTCTG	CTTCACTCTATT	1818
RESULT 34				
AAK04516				
ID AAK04516 standard; DNA; 1894 BP.				
AAK04516;				
05-NOV-2001 (first entry)				
Human brain expressed single exon probe SEQ ID NO: 4507.				
Human; brain expressed exon; gene expression analysis; probe; microarray;				
Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;				
88.				
Homo sapiens.				
WO200157275-A2.				
09-AUG-2001.				
30-JAN-2001; 2001WO-US000667.				
04-FEB-2000; 2000US-0180312P.				
26-MAY-2000; 2000US-0207456P.				
30-JUN-2000; 2000US-00608408.				
03-AUG-2000; 2000US-00612366.				
21-SEP-2000; 2000US-0214687P.				
27-SEP-2000; 2000US-0236359P.				
04-OCT-2000; 2000GB-00024263.				
(MOLE-) MOLECULAR DYNAMICS INC.				
Penn SG, Hanzel DK, Chen W, Rank DR;				
WPI; 2001-483446/52.				
Single exon nucleic acid probes for analyzing gene expression in human				
brains.				
Example 4: SEQ ID NO 4507; 650bp + Sequence Listing; English.				

Db	1801	CTCTGTCTTCACTCTATT	1818	Qy	121	AATTGAGATCGAATATATATAGTAGCAGAGGACCTTCAAAACACTGACACCTCGGGCCT	180
Db				Db	1081	AATTCAGAGTCCGATATATACGTAGAGCAAGGAGCTGCAAAACACTGAGACCTCGGGCCT	1140
Db				Qy	181	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	240
Db				Db	1141	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	1200
Qy				Qy	241	TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTTCTTCCAGAA	300
Db				Db	1201	GTACTCTCTTTGGACCCCTGTATCTTAAACCTCTTGTAAAGTTTGTCTTCTTCCAGAA	1260
Qy				Qy	301	TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCAGTAA	360
Db				Db	1261	CGAAGCAGTAAACTACAAATCGTTCTTCAATGAGGCCAGATGCGAGTCCAGTAA	1320
Qy				Qy	361	AATCTACCGTGGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATTTGA	420
Db				Db	1321	AATCTACCGACCGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATCAA	1380
Qy				Qy	421	AGTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	480
Db				Db	1381	AGGCACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	1440
Qy				Qy	481	AAGCAGTTAGAGCAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	540
Db				Db	1441	AAGCAGTTAGAGTGGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	1500
Qy				Qy	541	GGGTGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	600
Db				Db	1501	GGGGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	1560
Qy				Qy	601	ANCTGGGAAGTGACGGCATCCATCTTAAACATGGGCTTGCACCTTAGCTACACCCG	660
Db				Db	1561	AGCTGGGAAGTGACGGCTTCCACCTTTAAACACCGGGCTTGCACCTTAGCTACACCCG	1620
Qy				Qy	661	ACCAATC-----AGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGAGGATAA	711
Db				Db	1621	ACCAATCAGATCACTAAGAGAGCTCACTAATATGCTAATTAGGCAAAACAGAGGATAA	1680
Qy				Qy	712	GCAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAAGGATTGGGATATAA	771
Db				Db	1681	GAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAATGATCGGGATATAA	1740
Qy				Qy	772	ACTCAGGCAATCAAGCCACACACCCCTTGGGCTCCCTCCCATTTGATCGGAG	831
Db				Db	1741	ACCCAGGCAATCAGCCAGCTACAGCTACCCCTTTGGGCTCCCTCCCATTTGATCGGAG	1800
Qy				Qy	832	CTCTGTTTTCACCTCTATT	849
Db				Db	1801	CTCTGTTTTCACCTCTATT	1818
RESULT 36							
Id	AAI04422	standard; DNA; 1894 BP.					
Ac	AAI04422;						
Dt	09-OCT-2001	(first entry)					
De	Probe #4413	used to measure gene expression in human breast sample.					
Kw	Probe; human; breast disease; breast cancer; development disorder; ss;						
Os	inflammatory disease; proliferative breast disease; non-carcinoma tumour.						
Pn	Homo sapiens.						
Pd	WO200157270-A2.						
Pp	09-AUG-2001.						
	29-JAN-2001; 2001WO-US000661.						

Db	1801	CTCTGTCTTCACTCTATT	1818	Qy	121	AATTGAGATCGAATATATATAGTAGCAGAGGACCTTCAAAACACTGACACCTCGGGCCT	180
Db				Db	1081	AATTCAGAGTCCGATATATACGTAGAGCAAGGAGCTGCAAAACACTGAGACCTCGGGCCT	1140
Db				Qy	181	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	240
Db				Db	1141	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	1200
Qy				Qy	241	TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTTCTTCCAGAA	300
Db				Db	1201	GTACTCTCTTTGGACCCCTGTATCTTAAACCTCTTGTAAAGTTTGTCTTCTTCCAGAA	1260
Qy				Qy	301	TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCAGTAA	360
Db				Db	1261	CGAAGCAGTAAACTACAAATCGTTCTTCAATGAGGCCAGATGCGAGTCCAGTAA	1320
Qy				Qy	361	AATCTACCGTGGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATTTGA	420
Db				Db	1321	AATCTACCGACCGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATCAA	1380
Qy				Qy	421	AGTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	480
Db				Db	1381	AGGCACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	1440
Qy				Qy	481	AAGCAGTTAGAGCAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	540
Db				Db	1441	AAGCAGTTAGAGTGGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	1500
Qy				Qy	541	GGGTGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	600
Db				Db	1501	GGGGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	1560
Qy				Qy	601	ANCTGGGAAGTGACGGCATCCATCTTAAACATGGGCTTGCACCTTAGCTACACCCG	660
Db				Db	1561	AGCTGGGAAGTGACGGCTTCCACCTTTAAACACCGGGCTTGCACCTTAGCTACACCCG	1620
Qy				Qy	661	ACCAATC-----AGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGAGGATAA	711
Db				Db	1621	ACCAATCAGATCACTAAGAGAGCTCACTAATATGCTAATTAGGCAAAACAGAGGATAA	1680
Qy				Qy	712	GCAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAAGGATTGGGATATAA	771
Db				Db	1681	GAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAATGATCGGGATATAA	1740
Qy				Qy	772	ACTCAGGCAATCAAGCCACACACCCCTTGGGCTCCCTCCCATTTGATCGGAG	831
Db				Db	1741	ACCCAGGCAATCAGCCAGCTACAGCTACCCCTTTGGGCTCCCTCCCATTTGATCGGAG	1800
Qy				Qy	832	CTCTGTTTTCACCTCTATT	849
Db				Db	1801	CTCTGTTTTCACCTCTATT	1818
RESULT 36							
Id	AAI04422	standard; DNA; 1894 BP.					
Ac	AAI04422;						
Dt	09-OCT-2001	(first entry)					
De	Probe #4413	used to measure gene expression in human breast sample.					
Kw	Probe; human; breast disease; breast cancer; development disorder; ss;						
Os	inflammatory disease; proliferative breast disease; non-carcinoma tumour.						
Pn	Homo sapiens.						
Pd	WO200157270-A2.						
Pp	09-AUG-2001.						
	29-JAN-2001; 2001WO-US000661.						

Db	1801	CTCTGTCTTCACTCTATT	1818	Qy	121	AATTGAGATCGAATATATATAGTAGCAGAGGACCTTCAAAACACTGACACCTCGGGCCT	180
Db				Db	1081	AATTCAGAGTCCGATATATACGTAGAGCAAGGAGCTGCAAAACACTGAGACCTCGGGCCT	1140
Db				Qy	181	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	240
Db				Db	1141	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	1200
Qy				Qy	241	TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTTCTTCCAGAA	300
Db				Db	1201	GTACTCTCTTTGGACCCCTGTATCTTAAACCTCTTGTAAAGTTTGTCTTCTTCCAGAA	1260
Qy				Qy	301	TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCAGTAA	360
Db				Db	1261	CGAAGCAGTAAACTACAAATCGTTCTTCAATGAGGCCAGATGCGAGTCCAGTAA	1320
Qy				Qy	361	AATCTACCGTGGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATTTGA	420
Db				Db	1321	AATCTACCGACCGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATCAA	1380
Qy				Qy	421	AGTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	480
Db				Db	1381	AGGCACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	1440
Qy				Qy	481	AAGCAGTTAGAGCAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	540
Db				Db	1441	AAGCAGTTAGAGTGGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	1500
Qy				Qy	541	GGGTGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	600
Db				Db	1501	GGGGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	1560
Qy				Qy	601	ANCTGGGAAGTGACGGCATCCATCTTAAACATGGGCTTGCACCTTAGCTACACCCG	660
Db				Db	1561	AGCTGGGAAGTGACGGCTTCCACCTTTAAACACCGGGCTTGCACCTTAGCTACACCCG	1620
Qy				Qy	661	ACCAATC-----AGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGAGGATAA	711
Db				Db	1621	ACCAATCAGATCACTAAGAGAGCTCACTAATATGCTAATTAGGCAAAACAGAGGATAA	1680
Qy				Qy	712	GCAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAAGGATTGGGATATAA	771
Db				Db	1681	GAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAATGATCGGGATATAA	1740
Qy				Qy	772	ACTCAGGCAATCAAGCCACACACCCCTTGGGCTCCCTCCCATTTGATCGGAG	831
Db				Db	1741	ACCCAGGCAATCAGCCAGCTACAGCTACCCCTTTGGGCTCCCTCCCATTTGATCGGAG	1800
Qy				Qy	832	CTCTGTTTTCACCTCTATT	849
Db				Db	1801	CTCTGTTTTCACCTCTATT	1818
RESULT 36							
Id	AAI04422	standard; DNA; 1894 BP.					
Ac	AAI04422;						
Dt	09-OCT-2001	(first entry)					
De	Probe #4413	used to measure gene expression in human breast sample.					
Kw	Probe; human; breast disease; breast cancer; development disorder; ss;						
Os	inflammatory disease; proliferative breast disease; non-carcinoma tumour.						
Pn	Homo sapiens.						
Pd	WO200157270-A2.						
Pp	09-AUG-2001.						
	29-JAN-2001; 2001WO-US000661.						

Db	1801	CTCTGTCTTCACTCTATT	1818	Qy	121	AATTGAGATCGAATATATATAGTAGCAGAGGACCTTCAAAACACTGACACCTCGGGCCT	180
Db				Db	1081	AATTCAGAGTCCGATATATACGTAGAGCAAGGAGCTGCAAAACACTGAGACCTCGGGCCT	1140
Db				Qy	181	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	240
Db				Db	1141	CCTCAGCAATGAGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATATATT	1200
Qy				Qy	241	TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTTCTTCCAGAA	300
Db				Db	1201	GTACTCTCTTTGGACCCCTGTATCTTAAACCTCTTGTAAAGTTTGTCTTCTTCCAGAA	1260
Qy				Qy	301	TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCAGTAA	360
Db				Db	1261	CGAAGCAGTAAACTACAAATCGTTCTTCAATGAGGCCAGATGCGAGTCCAGTAA	1320
Qy				Qy	361	AATCTACCGTGGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATTTGA	420
Db				Db	1321	AATCTACCGACCGACCCCTGGACCGCTGCTAGATATGCTCTGATTTAATGACATCAA	1380
Qy				Qy	421	AGTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	480
Db				Db	1381	AGGCACCCCTCCGAGGAAATCTCAACTGCAACCCCTTACTACACTCCCAATTCAGTAG	1440
Qy				Qy	481	AAGCAGTTAGAGCAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	540
Db				Db	1441	AAGCAGTTAGAGTGGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	1500
Qy				Qy	541	GGGTGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	600
Db				Db	1501	GGGGGACTAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	1560
Qy				Qy	601	ANCTGGGAAGTGACGGCATCCATCTTAAACATGGGCTTGCACCTTAGCTACACCCG	660
Db				Db	1561	AGCTGGGAAGTGACGGCTTCCACCTTTAAACACCGGGCTTGCACCTTAGCTACACCCG	1620
Qy				Qy	661	ACCAATC-----AGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGAGGATAA	711
Db				Db	1621	ACCAATCAGATCACTAAGAGAGCTCACTAATATGCTAATTAGGCAAAACAGAGGATAA	1680
Qy				Qy	712	GCAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAAGGATTGGGATATAA	771
Db				Db	1681	GAATAGCCAATCATCTATTGCTGAGAGCAGCGGGGAAGACAAATGATCGGGATATAA	1740
Qy				Qy	772	ACTCAGGCAATCAAGCCACACACCCCTTGGGCTCCCTCCCATTTGATCGGAG	831
Db				Db	1741	ACCCAGGCAATCAGCCAGCTACAGCTACCCCTTTGGGCTCCCTCCCATTTGATCGGAG	1800
Qy				Qy	832	CTCTGTTTTCACCTCTATT	849
Db				Db	1801	CTCTGTTTTCACCTCTATT	1818
RESULT 36							
Id	AAI04422	standard; DNA; 1894 BP.					
Ac	AAI04422;						
Dt	09-OCT-2001	(first entry)					
De	Probe #4413	used to measure gene expression in human breast sample.					
Kw	Probe; human; breast disease; breast cancer; development disorder; ss;						
Os	inflammatory disease; proliferative breast disease; non-carcinoma tumour.						
Pn	Homo sapiens.						
Pd	WO200157270-A2.						
Pp	09-AUG-2001.						
	29-JAN-2001; 2001WO-US000661.						

Db	1801	CTCTGTCTTCACTCTATT	1818																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		</
----	------	--------------------	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

XX PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX DR Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX Claim 25; SEQ ID NO 4413; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridizes at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match 53.0%; Score 704.8; DB 5; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAACCAGCGAAAGAGGGGAACTGTTATATTTT 60
Db 961 TCAAAATCGAAGAGCTTTAGACTTCTAACCAGCGAAAGAGGGGAACTGTTATATTTT 1020
QY 61 AGGGAAGAACTCTGTAGTATGTTAACTAATCTCGAATCATCTAGAGAAAGTTAAAGA 120
Db 1021 AGGGAAGAACTCTGTATGTTATGTTAACTAATCTCGAATCATCCGGAAGTTAAAGA 1080
QY 121 AATTTGAGATCGAATATAATATAGACAGAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 1081 AATTCAGGTCGAATATAACGTAGAGCAAGAGGCTGCAGAAACACTGGACCTGGGGCT 1140
QY 181 CCTCAGCAATGGATGCCCTGAGCTCTCCCTTCTTAGACCTCTAGCAGCTATAATTT 240
Db 1141 CCTCAGCAATGGATGCCCTGAGTCTCCCTTCTTAGACCTCTAGCAGCTATAATTT 1200
QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTCTCTGTTAGTTTCTCTTCCAGAT 300
Db 1201 GTTACTCTCTTTGGACCTGATCTTAACTCTCTGTTAGTTTCTCTTCCAGAT 1260
QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTTCAATGGAACCCAGATGCAATCACTAA 360
Db 1261 CGAAGCAGTAAACTACAAATCGTTCTTCAATGGAACCCAGATGCAATCACTAA 1320
QY 361 AATCTACCTGACCCCTGGACCGGCTCTAGATGCTGATGTTATGACATTGA 420
Db 1321 AATCTACCGACCCCTGGACCGGCTCTAGCCATGCTGATGTTATGACATCAA 1380
QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGACAAACCCCTACTACACTCCAATTCAGTAGG 480
Db 1381 AGGCACCCCTCCCGAGGAATCTCAACTGACAAACCTCTACTACGCCCAATTCAGCAGG 1440
QY 481 AAGCAGTTAGACAGTTGTGACGCAACCTTCCCAACAGTACTTGGGTTTCTGTTGAGA 540

Db 1441 AAGCAGTTAGAGTGGTGGTGGCCAACTCCCAACAGCAGTGGGTTTCTCTGTTGAGA 1500
QY 541 GGGTGGACTGAGAGACGAGTACTAGTGGATTTCTTAGCTGACTAAGAATCCCAAGCCT 600
Db 1501 GGGGGGACTGAGAGACGAGGAATAACTAGATTTCTAGACCAACTAAGAATCCCTAAGACT 1560
QY 601 ANCTGGGAAGTGCACCGCATCCATCTTTAAACATGGGGCTTGGCAACTTAGCTCACACCCG 660
Db 1561 AGCTGGGAAGTGCACCGCTTCCACCTTTAAACACCGGGCTTGAACCTTAGCTCAGGCCA 1620
QY 661 ACCAATC-----AGAGAGTCTACTAAATGCTAATCAGGCAAAAACAGGAGGTAAA 711
Db 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTAATTAGGCAAAAACAGGAGATAA 1680
QY 712 GCATAGCCCAATCATCTATTGCTGAGACGACGCGGAGGAGCAAGGATGGGATATAA 771
Db 1681 GAAATAGCCCAATCATCTGTTGCTGACAGCAGCAGGAGGAGCAATGATCGGATATAA 1740
QY 772 ACTCAGGCATTTCAAGCCAGCAACAGCAACCCCTTTTGGGTCCCTCCCATTTGTATGGAG 831
Db 1741 ACCAGGCATTCGAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGTATGGAG 1800
QY 832 CTCTGTTTTCACCTTATT 849
Db 1801 CTCTGTTTTCACCTTATT 1818
RESULT 37
ABS04589
ID ABS04589 standard; DNA; 1894 BP.
AC ABS04589;
XX
DT 19-AUG-2002 (first entry)
XX Human genome-derived single exon probe from lung SEQ ID NO 4580.
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX Homo sapiens.
OS
XX
XX WO200186003-A2.
PN
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US0000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2002-114183/15.
DR
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
PT
XX Claim 1; SEQ ID NO 4580; 634pp; English.
PS

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Herxsky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 53.0%; Score 704.8; DB 6; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTATTTT 60
DB 961 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCGGAAAGCGGGGAACCTTTTATTTT 1020
QY 61 AGGGGAAGATGCTGTTAGTATGTTTAACTCAATCTGGAATCATCTGAGAAAGTTAAAGA 120
DB 1021 AGAGGAAAAATGCTGTTGTTATGTTTAACTCAATCCGAATCATCACCAGAAAGTTAAAGA 1080
QY 121 AATTGAGATCGAATATATGATAGACAGAGACCTTCAAAACACCTGCGGCGCT 180
DB 1081 AATTCAAGTCGAATATATGATAGCAAGAGAGCTGCAAAACACCTGCGGCGCT 1140
QY 181 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 240
DB 1141 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 1200
QY 241 TTTACTGCTTTTGGACCTGATATCTTCAACTCTCTTTGTTAAGTTGTCCTTCCAGAT 300
DB 1201 GTTACTCTCTTTGGACCTGATATCTTCAACTCTCTTTGTTAAGTTGTCCTTCCAGAT 1260
QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
DB 1261 CGAAGCAGTAAACTACAAATCGTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 1320
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
DB 1321 AATCTACCGGACCCCTGGACCGGCTGCTAGCCATGCTCTGATGTTAATGACATCAA 1380

QY 421 AGTCACCCCTCCGAGGAAATCTCACTGCAACACCCCTACTACTCAATTTAGTAGG 480
DB 1381 AGGACCCCTCCGAGGAAATCTCACTGCAACACCCCTACTACTAGCCCAATTTAGCAGG 1440
QY 481 AAGCAGTTAGAGCAGTGTCTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
DB 1441 AAGCAGTTAGAGTGTGTTGTTGCGCAACCTCCCAACAGCAGTTGGGTTTTCTGTTGAGA 1500
QY 541 GGGTGGACTGAGACAGAGCTAGCTGATTTCTTAGCTGACTAAGAAATCCNAAAGCCT 600
DB 1501 GGGGGGACTGAGACAGAGATACTAGATTTCTTAGCCAACTAAGAAATCCCTAAGACT 1560
QY 601 ANCTGGGAAGGTGACCGCATCTTTTAAACATGGGCTTGCAACTTAGCTTCACACCCG 660
DB 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTTCACGCCA 1620
QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGAGTAA 711
DB 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTTAATAGGCAAAACAGAGATATA 1680
QY 712 GCNATAGCCATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTCGGATATAA 771
DB 1681 GAAATAGCCATCATCTGTTGCTGACAGCAGCAGGAGGACAAATGATCGGATATAA 1740
QY 772 ACTCAGGCAATTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCCAATTTGATGGGAG 831
DB 1741 ACCAGGCAATTCAGCCAGCTACAGTACCTCTTTGGGTCCCTCCCTTTGATGGGAG 1800
QY 832 CTCTGTTTCACTCTATT 849
DB 1801 CTCTGTTTCACTCTATT 1818
RESULT 38
AAS77313
ID AAS77313 standard; cDNA; 2942 BP.
XX AAS77313;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #13117.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG13126.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 13117; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

1982	TGTCACAGACCCACAGCTGACTCTCCATCCCTCTG-----	2015	-----
1006	CTGATCCAGCGAGGTACCCATTGCCACTCCCGATCAGGCTAAAGGCTTGCCATGTGTCCT	1065	CTGATCCAGCGAGGTACCCATTGCCACTCCCGATCAGGCTAAAGGCTTGCCATGTGTCCT
2016	--GATCCAGCGAGGCGCCCATTCGTCTCTGATCGGGCTAAAGGCTTGCCCTGTGTCCT	2073	--GATCCAGCGAGGCGCCCATTCGTCTCTGATCGGGCTAAAGGCTTGCCCTGTGTCCT
1066	GCATGGCTAAGTG-CTTGGGTTTGTCTTAATAGAACTGAACACTGTGTCACTGGGTTCCAT	1124	GCATGGCTAAGTG-CTTGGGTTTGTCTTAATAGAACTGAACACTGTGTCACTGGGTTCCAT
2074	GCAGGGCTAAGTGCCCCAGGATTGGTCTTAATCTAGCTAAACACTGTAGTCTCTGGGTTCCAC	2133	GCAGGGCTAAGTGCCCCAGGATTGGTCTTAATCTAGCTAAACACTGTAGTCTCTGGGTTCCAC
1125	GGTTCTCTTCCATGACCCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCCAAG	1184	GGTTCTCTTCCATGACCCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCCAAG
2134	GAATCTCTTCCATGATGCCACAGCTTCTAATAGAGCTATAACACTCACTGCATGGCCCCACA	2193	GAATCTCTTCCATGATGCCACAGCTTCTAATAGAGCTATAACACTCACTGCATGGCCCCACA
1185	ATTCCATTCTTGATATCTGTGAGGCCAAGAA-CCCCAGGCTCAGAGAAANGTGAGGTTGC	1243	ATTCCATTCTTGATATCTGTGAGGCCAAGAA-CCCCAGGCTCAGAGAAANGTGAGGTTGC
2194	GTTCCATTCTTTGGAAATCCGTGAGGCCAAGAACCCCCAGGTCACAGAAATGAGGCTTGC	2253	GTTCCATTCTTTGGAAATCCGTGAGGCCAAGAACCCCCAGGTCACAGAAATGAGGCTTGC
1244	CACCAATTGGGAAGTGGCCCACTGCCAATTTTGTAGCGGCCCAACCAACATCTTTGGGAGCT	1303	CACCAATTGGGAAGTGGCCCACTGCCAATTTTGTAGCGGCCCAACCAACATCTTTGGGAGCT
2254	CACCATCTCGGNAAGAGSCCGTCACCATCTTGGAAAGCGACTGCCACCATCTTTGGGAAC	2313	CACCATCTCGGNAAGAGSCCGTCACCATCTTGGAAAGCGACTGCCACCATCTTTGGGAAC
1304	GTGGGACGAAGGATCCCC	1321	GTGGGACGAAGGATCCCC
2314	TGGGGACGAAGGACCCCC	2331	TGGGGACGAAGGACCCCC

RESULT 39	
AA565964	
ID	AA565964 standard; cDNA; 893 BP.
AC	AA565964;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #1768.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Dmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	P-PSDB; ARG01777.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 1768; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 893 BP; 223 A; 264 C; 202 G; 204 T; 0 U; 0 Other;
 Query Match 49.0%; Score 651; DB 5; Length 893;
 Best Local Similarity 85.0%; Pred. No. 1.1e-202;
 Matches 757; Conservative 0; Mismatches 117; Indels 17; Gaps 2;
 321 TAGTCTTCAATGGAACCCAGATGCGATGCTCAATCAATCTACCGTGACCCCTGG 380
 Db |||||
 10 TAAAACTACAAATGGAGCCAAAGATGCGATGCTCAATCAATCTACCGTGACCCCTGG 69
 QY |||||
 381 ACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTAAGTCAACCTCCCGAGGAAA 440
 Db |||||
 70 ACCGGCTGCTAGCCACGATCTGATGTTAATGACATTAAGTCAACCTCCCGAGGAAA 129
 QY |||||
 441 TCTCAACTGCACAAACCCCTACTACACTTCAATTCAGTAGGAGGAGTGTAGACAGTTGTC 500
 Db |||||
 130 TCTAGCTGCACAACTCTACTACGCCCAATTCAGGAGGAGGAGTGTAGACAGTTGTC 189
 QY |||||
 501 AGCCAACTCCCAACAGTACTGTTGGTCTCTGTTGAGAGGTTGAGTGCAGACAGGA 560
 Db |||||
 190 GGCCAACTCCCAACAGTACTGTTGGTCTCTGTTGAGAGGTTGAGTGCAGACAGGA 249
 QY |||||
 561 CTAGCTGGATTTCTTAGGCTGCTAGAGTCCCAAGCTTACTGGAAGGTTGACCGCAT 620
 Db |||||
 250 CTAGCTGGATTTCTTAGGCTGCTAGAGTCCCAAGCTTACTGGAAGGTTGACCGCAT 309
 QY |||||
 621 CCATCTTTAAACATGGGGCTTGCACTTAGCTCACAACCCGACCAATC-----AGAG 671
 Db |||||
 310 CCACCTTTAAACACGGTGTCTGCACTAGCTCACAACCCGACCCATCAGGTAAGAAAGAG 369
 QY |||||
 672 AGCTACTAAATGCTTAATCAGGCAAAAACAGAGGTAAGCAATAGCAATCACTATT 731
 Db |||||
 370 AGCCCGCTAAATGCTTAATTAGGCAAAAACAGAGGTAAGCAATAGCAATCACTATT 429
 QY |||||
 732 GCTGAGACACAGCGGAGGACAGGATGCGATATAAACTCAGGCAATTCAGGCGAGC 791
 Db |||||
 430 GCTGAGACACAGCGGAGGACAGATGATGAGTATATAACCCAGGCAATTCAGGCGAGC 489
 QY |||||
 792 AACAGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTTCACTTATTTC 851
 Db |||||
 490 AAGCACTACCTTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTTCACTTATTAA 549
 QY |||||
 852 ACTCTATTAAATCATGCAACTGCACTCTTCTGCTCGTGTGTTTTTATGCTCAAGCTGAG 911
 Db |||||
 550 ATCTTGCAACTGCA-----CACTCTTCTGCTGTGTTTTGTTACGGCTCGAGCTGAG 601
 QY |||||
 912 CTTTCTGCTGCCATCCACCACTGCTGTTGGCAACCGCTCAGACCCGCTGCTGACTTCCA 971
 Db |||||
 602 CTTTCTGCTCACCCTCACCACACTGCTGTTTGGCAACCGCTGCTGAGACCCATCGCTGACTTCCA 661
 QY |||||
 972 TCCTTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCGAGGTACCCATTGCCA 1031
 Db |||||
 662 CCCTCCAGATAGGCGAGGTTGCGCTGCTGTTCTGATATAGCAAGGCGCCATTGGCG 721
 QY |||||
 1032 CTCCTGATCAGGCTAAGGCTTGGCAATGTTGTTCTGATGCTGCTGCTGCTGCTGCTGCTC 1091
 Db |||||
 722 CTCCTCAATCGGCTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 781
 QY |||||
 1092 TAATAGAACTGAACACTGGTCACTGGGTTCCATGTTCTCTTCCATGACCCCGGCTTCT 1151

Db |||||
 782 TAATCCAGCTGAACACTAGTGTGCTGGTTCCACGGTTCTCTTCCATGACCCACGGCTTCT 841
 QY |||||
 1152 AATAGCTATAACACTCAACCGCATGCGCCAGATTCCTTCCATTCCTTGGTATC 1202
 Db |||||
 842 AATAGAGCTATAACACTCAACCGCATGCGCCAGGTTCCATTCCTTGGATC 892
 RESULT 40
 ABN97978
 ID ABN97978 standard; DNA; 46340 BP.
 XX AC ABN97978;
 XX DT 01-AUG-2002 (first entry)
 XX DE Human retroviral sequence HI3.
 XX KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
 XX OS Human endogenous retrovirus.
 XX PN WO9967395-A1.
 XX PD 29-DEC-1999.
 XX PF 23-JUN-1999; 99WO-FR001513.
 XX PR 23-JUN-1998; 98FR-00007920.
 XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX PI Alliel PM, Perin J, Rieger F;
 XX DR WPI; 2000-160587/14.
 XX PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
 PT for diagnosis, treatment and prevention of autoimmune and neurological
 PT diseases.
 PS Claim 15; Page 186-199; 225pp; French.
 CC The present invention relates to new nucleic acid sequences of human
 CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
 CC Regulatory elements associated with HERV-7q may alter expression of other
 CC genes (even remote genes) on the same chromosome, inducing immunological
 CC and/or neurological changes (which may be pathological or protective/
 CC curative). HERV-7q peptides can be used to improve efficiency of the
 CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
 CC sequences can be used in immunogenic or vaccinating compositions, for
 CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention
 XX SQ Sequence 46340 BP; 16104 A; 8738 C; 8434 G; 13064 T; 0 U; 0 Other;
 Query Match 48.4%; Score 643.4; DB 3; Length 46340;
 Best Local Similarity 85.4%; Pred. No. 3.6e-199;
 Matches 781; Conservative 0; Mismatches 88; Indels 46; Gaps 4;
 1 TCATAATCGAAGCTTTAGACTTCTACCCCAAAAGAGGGGAACTGTTTATTTT 60
 Db |||||
 32993 TCATAATCGAAGCTTTAGACTTCTACCTCTGAAGAGGGGAACTGTTTATTTT 33052
 QY |||||
 61 AGCGGAAGAAATGCTGTAG-----TATGTTAATCA 90
 Db |||||
 33053 AGCGGAAGAAATGTTGTTATTTTATGCGGAGAAATGTTGTTATTTAATCA 33112
 QY |||||
 91 ATCTGGAATCATTAATGAGAAAGTTAAAGAAATTTGAGATCGAATATATATGATGACGAGA 150
 QY |||||

Fri Feb 25 16:26:28 2005

us-09-319-156b-12.rng

Db 33113 ATCCTGAATTCTCACAGAAAAGTTGAAGAAATTCGAGATTGAATACAACTAGAACAGA 33172
Qy 151 GGACCTTC-AAAACACTGCACCCCTGGGGCTCTCCTCAGCCAATGGATGCCCTGGACTCTCC 209
Db 33173 GGAGCTTCAAAAACACAGACCCCTGGGGCTCTCCTCAGCCAATGGATGCCCTGGATCTCC 33232
Qy 210 CCTTCTTAGGACCTCTAGCAGCTATAATATTTTACTCCTCTTTGGACCCCTGTATCTTCA 269
Db 33233 CCTTCTTAGGATCTCTAGCAGCTCTAATATTTGATACTCTCTTTGGACCCCTGTATCTTTA 33292
Qy 270 ACTTCCTTTGTTAAGTTGTCTCTTCCAGAAATGAAGCTGTAAAGCTACAAATAGTTCTTTC 329
Db 33293 ACCTCCTTTGTTAAGTTGTCTCTTCCAGAAATCNAAGTTGTAAAGCTACAAATCGTTCTTC 33352
Qy 330 AAATGGRAACCCAGATGCAGTCCATGACTAAAATCTACCGTGGACCCCTGGACCCGGCTG 389
Db 33353 AAATGGAAACCCAGATGAAGTCCATGACTAGATCTACCGTGGACCCCTGGACCCGGCTA 33412
Qy 390 CTAGACTATGCTCTGAATGTTAATGACATTGAAGTCAACCCCTCCCGAGGAAATCTCAACTG 449
Db 33413 CTAGCCCATGCTCCAAATTTGAATGATATCGAAACGACCCCTCCCGAGGAAATCTCAACTG 33472
Qy 450 CACAACCCCTACTACACTCCAAATTCAGTAGGAGCAGTTAGAGCAGTTGTCTCAGCCAACT 509
Db 33473 CACAACCCCTACTATGCCCCCAATCCCGAGGAAGCAGTTAGACTGTCGTCTCAGCCAACT 33532
Qy 510 CCCCAACAGTACTTGGGTTTCTGTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGA 569
Db 33533 CCCCAACAGCACTTGGGTTTCTGTTGAGTGGGGGACTGAGAGACAGGATTAGCTGGA 33592
Qy 570 TTTCCTAGGCTGACTAAGAAATCCCNVAGCCTTANCTGGGAAGTGACCGCATCCATCTTTA 629
Db 33593 TTTCCTAGGCCGACTAAGAAATCCCAAGCCTAGCTGGGAAGTGACCCACATCCACCTTTA 33652
Qy 630 AACATGGGGCTTGGCAACTTAGCTCACCCGACCAATC-----AGAGAGCTCACTA 680
Db 33653 AACATGGGGCTTGGCAACTTAGCTCACCCGACCAATCAGGTAGTAAAGAGAGCTCACTA 33712
Qy 681 AAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCCAATCATCTATTGCTGAGAG 740
Db 33713 AAATGCTAATTAGACAAAAACAGGAGGTAAAAAATAGCCAATCATCTATCGCTGAGAG 33772
Qy 741 CACAGCGGAGGACAGAGATTGGGATATAAACTCAGGCATTCAAGCCAGCAACAGCAAC 800
Db 33773 CACAGCGGAGGACAGATATCGGGATATAAAACCCAGGCATTCAAGCCGCAACGCTAC 33832
Qy 801 CCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTCACTCTATTTCACTCTATTA 860
Db 33833 CTTCCTTGGGTCCCTCCCTCTTTGATGGGAGCTCTCTCT-----GTCTTCACTCTATTA 33886
Qy 861 AATCATGCAACTGCA 875
Db 33887 AATATTGCAACTGCA 33901

Search completed: February 21, 2005, 07:46:02
Job time : 696.525 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:54:34 ; Search time 234.166 Seconds
(without alignment)
9286.612 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatcgagagcttag.....gcaaggatccccagtaaca 1329

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1326	99.8	1329	4	US-08-979-847B-108
2	1035.2	77.9	13537	4	Sequence 17382, A
3	731.8	55.1	2946	3	Sequence 3, Appli
4	729.2	54.9	2763	4	Sequence 5640, Ap
5	728.4	54.8	168394	4	Sequence 13002, A
6	654.8	49.3	77772	4	Sequence 17417, A
7	654.8	49.3	77772	4	Sequence 12249, A
8	643.4	48.4	145320	4	Sequence 15858, A
9	637.8	48.0	780	4	Sequence 385, App
10	634.6	47.8	245286	4	Sequence 15497, A
11	582	43.8	219964	4	Sequence 15086, A
12	565.6	42.6	84571	4	Sequence 17420, A
13	562	42.3	194790	4	Sequence 15393, A
14	545	41.0	635	4	Sequence 102, App
15	500.2	37.6	140234	4	Sequence 17002, A
16	478.6	36.0	116552	4	Sequence 13413, A
17	468.6	35.3	131631	4	Sequence 11757, A
18	467.2	35.2	89584	4	Sequence 17068, A
19	446	33.6	154023	4	Sequence 17057, A
20	445.2	33.5	251672	4	Sequence 17296, A
21	445.2	33.5	251682	4	Sequence 11973, A
22	440.4	33.1	85523	4	Sequence 21, Appl
23	435.8	32.8	99580	4	Sequence 17411, A
24	434.8	32.7	601	4	Sequence 202370, A
25	427.8	32.2	601	4	Sequence 44299, A
26	411.4	31.0	57507	4	Sequence 15019, A
27	399.8	30.1	39686	4	Sequence 13633, A

C 28	399.8	30.1	49487	4	US-09-949-016-11770	Sequence 11770, A
C 29	387.4	29.1	601	4	US-09-949-016-44300	Sequence 44300, A
C 30	371	27.9	601	4	US-09-949-016-195793	Sequence 195793, A
C 31	369.6	27.8	601	4	US-09-949-016-44298	Sequence 44298, A
C 32	345.4	26.0	601	4	US-09-949-016-146564	Sequence 146564, A
C 33	344.6	25.9	443	4	US-09-513-999C-21394	Sequence 21394, A
C 34	343.6	25.9	149971	4	US-09-949-016-13590	Sequence 13590, A
C 35	338.8	25.5	601	4	US-09-949-016-133739	Sequence 133739, A
C 36	336	25.3	601	4	US-09-949-016-120277	Sequence 120277, A
C 37	324.6	24.4	601	4	US-09-949-016-202257	Sequence 202257, A
C 38	317.8	23.9	601	4	US-09-949-016-146565	Sequence 146565, A
C 39	317.6	23.9	601	4	US-09-949-016-44297	Sequence 44297, A
C 40	316.4	23.8	601	4	US-09-949-016-44296	Sequence 44296, A
C 41	311.6	23.4	601	4	US-09-949-016-195792	Sequence 195792, A
C 42	305.4	23.0	601	4	US-09-949-016-183390	Sequence 183390, A
C 43	302	22.7	1481	4	US-08-979-847B-105	Sequence 105, App
C 44	273.6	20.6	601	4	US-09-949-016-195794	Sequence 195794, A
C 45	268.4	20.2	601	4	US-09-949-016-44310	Sequence 44310, A

ALIGNMENTS

RESULT 1
US-08-979-847B-108
; Sequence 108, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKER, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1329 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-08-979-847B-108

Query Match									
Best Local Similarity 99.8%; Score 1326; DB 4; Length 1329;									
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTTATTTT	60						
Db	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTTATTTT	60						
Qy	61	AGGGGAAGATGCTGTGTAGTATGTAAATCAATCTGGAATCAATTAAGTGAAGGTTAAAGA	120						
Db	61	AGGGGAAGATGCTGTGTAGTATGTAAATCAATCTGGAATCAATTAAGTGAAGGTTAAAGA	120						
Qy	121	AATTGAGATCGAATATAATATAGAGCAGAGGACCTTCAAAACAATGCACTGCAACCTGGGCGCT	180						
Db	121	AATTGAGATCGAATATAATATAGAGCAGAGGACCTTCAAAACAATGCACTGCAACCTGGGCGCT	180						
Qy	181	CCTCAGCCAAATGGATGCCCTGGACTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT	240						
Db	181	CCTCAGCCAAATGGATGCCCTGGACTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT	240						
Qy	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTCTGTTAAAGTTTGTCTCTCCAGAAT	300						
Db	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTCTGTTAAAGTTTGTCTCTCCAGAAT	300						
Qy	301	TGAAGCTGAAGCTACAAATAGTTTCTTCAATGGAACCCAGATGCAATGCACTAA	360						
Db	301	TGAAGCTGAAGCTACAAATAGTTTCTTCAATGGAACCCAGATGCAATGCACTAA	360						
Qy	361	AATCTACCGTGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAATGACATTGA	420						
Db	361	AATCTACCGTGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAATGACATTGA	420						
Qy	421	AGTCAACCCCTCCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAATCAGTAGG	480						
Db	421	AGTCAACCCCTCCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAATCAGTAGG	480						
Qy	481	AAGCAGTTAGAGCAGTTGTGACGCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540						
Db	481	AAGCAGTTAGAGCAGTTGTGACGCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540						
Qy	541	GGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACATAAGAAATCCNAAGCCT	600						
Db	541	GGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACATAAGAAATCCNAAGCCT	600						
Qy	601	ANCTGGGAAGTGACCGCATCACTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660						
Db	601	ANCTGGGAAGTGACCGCATCACTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660						
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720						
Db	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720						
Qy	721	AATCATCTATTCCCTGAGAGCACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780						
Db	721	AATCATCTATTCCCTGAGAGCACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780						
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTTT	840						
Db	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTTT	840						
Qy	841	CACCTATTTCACCTATTAAATCATGCAACTGCACTCTCTGTCGGTCCGTTTTTATGG	900						
Db	841	CACCTATTTCACCTATTAAATCATGCAACTGCACTCTCTGTCGGTCCGTTTTTATGG	900						
Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCCACTGCTGTTTGGCCACCGTCACAGACCCGCT	960						
Db	901	CTCAAGCTGAGCTTTTGTTCGCCATCCCACTGCTGTTTGGCCACCGTCACAGACCCGCT	960						
Qy	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGGAGGT	1020						
Db	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGGAGGT	1020						

RESULT 2
US-09-949-016-17382
; Sequence 17382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17382
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17382

Query Match									
Best Local Similarity 77.9%; Score 1035.2; DB 4; Length 13537;									
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;									
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTTATTTT	60						
Db	10675	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTT	10734						
Qy	61	AGGGGAAGATGCTGTGTAGTATGTAAATCAATCTGGAATCAATTAAGGTTAAAGA	120						
Db	10735	AGGGGAAGATGCTGTGTATATGTTATCAATCAATCGGAATCTGCTACTGAGAAGTTAAAGA	10794						
Qy	121	AATTGAGATCGAATATAATATAGAGCAGAGGACCTTCAAAACAATGCACTGCAACCTGGGCGCT	180						
Db	10795	AATTGAGATCGAATATAAATGAGAGGAGGAGCTTCGAAACACTGGACCTGGGCGCT	10854						
Qy	181	CCTCAGCCAAATGGATGCCCTGGACTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT	240						
Db	10855	CCTCAGCCAAATGGATGCCCTGGACTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT	10914						
Qy	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTCTGTTAAAGTTTGTCTCTCCAGAAT	300						
Db	10915	GCTACTCTCTTTGGACCTGTATCTTCAACTCTCTGTTAACTCTTGTCTCTCCAGAAT	10974						

```

QY 301 TGAAGCTGAAGAGTACAAATAGTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 360
Db 10975 CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATGCGAGTCCAAAGACTAA 11022
QY 361 AATCTACCGTGGACCCCTGGAGCCGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 11023 GATCTACCGCAGACCCCTGGAGCCGCTGCTAGCCACGATCTGATGTTAATGACATCA 11082
QY 421 AGTCACCCCTCCCGAGGAATCTCAATGTCACAAACCCCTACTACACTCAATTCAGTAGG 480
Db 11083 AGGCACCCCTCTCTAGGAAATCTCAGTGTGACAACTCTACTACGCCCAATTCAGCAG 11142
QY 481 AAGCAGTTAGACAGTTGTGACCAACCTCCCAACAGTACTTGGTTCCTGTTGAGA 540
Db 11143 AAGCAGTTAGACAGTGTGTCGCCCAACCTCCCAACAGCAGTACTTGGTTCCTGTTGAGA 11202
QY 541 GGTGGAGTGTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCAAGCCT 600
Db 11203 TGGGGACTGTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCAAGCCT 11262
QY 601 ANCTGGGAAGTGTGACCGCATCCATCTTTAAACATGCGGCTTGCAACTTTAGCTCACACCG 660
Db 11263 AGCTGGGAAGTGTGACCACTCACCTTTAAACATGCGGCTTGCAACTTTAGCTCACACCTG 11322
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
Db 11323 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 11382
QY 721 AATCATCTAATTCCTGTAGAGACAGCGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
Db 11383 AATCATCTAATTCCTGTAGAGACAGCGGAAGCAAGGATTTGGGATATAAACTCAGGCA 11442
QY 781 TTCAAGCCAGCAACCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTTT 840
Db 11443 TTCAAGCCAGCAACCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTGTTTT 11502
QY 841 CACTCTAATTTCACTTAAATCATGCACTGCACTCTTCTGCTCGGTTTATGG 900
Db 11503 CATGCTAATTTCACTTAAATCATGCACTGCACTCTTCTGCTCGGTTTATGG 11562
QY 901 CTCAGCTGAGCTTTTGTTCGCATCCACACTGCTGTTTGGCCAGCTCACAGACCCGCT 960
Db 11563 CTTGAGCTGAGCTTTTGTTCGCATCCACACTGCTGTTTGGCCAGCTCACAGACCCGCT 11622
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGTATCCAGCAGGT 1020
Db 11623 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGTATCCAGCAGGT 11682
QY 1021 ACCATTGCACTCCCGATCAGGCTAAGGCTTGCATTTGCTTCTGATGGCTTAAGTGC 1080
Db 11683 ACCATTGCGGCTCCCAATCGGCTAAGGCTTGCATTTGCTTCTGATGGCTTAAGTGC 11742
QY 1081 TGGGTTTGTCTAATAGAACTGAACACTGGTCTGCTGGGTTCCATGTTCTTCCATGAC 1140
Db 11743 TGGGTTTGTCTAATAGAACTGAACACTGGTCTGCTGGGTTCCATGTTCTTCTGATGAC 11802
QY 1141 CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA 1200
Db 11803 CCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTT-GAA 11861
QY 1201 TCTGTAGGCAAGAACCCAGTGTAGAGANGTGGGCTTGCACCATTTGGGAAG 1257
Db 11862 TCCATAAGGCCAAGAACCCAGGTCAGAGAACGAGGCTTGCACCATTTGGGAAG 11918

```

RESULT 3
 US-09-175-928-3
 ; Sequence 3, Application US/09175928A
 ; Patent No. 6312921
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.

; APPLICANT: LaVallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Mi, Sha
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6006B AJ172A
 ; CURRENT APPLICATION NUMBER: US/09/175,928A
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2946
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-175-928-3

Query Match 55.1%; Score 731.8; DB 3; Length 2946;
 Best Local Similarity 90.7%; Pred. No. 9, 1e-240;
 Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACTGTTATTTT 60
 Db 2067 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 2126
 QY 61 AGGGAAGAGTCTCTTAGTATGTTAAATCAATCTGGAATCATTTACTTGAGAAAAGTTAAAGA 120
 Db 2127 AGGGAAGAGTCTCTTAGTATGTTAAATCAATCTGGAATCATTTACTTGAGAAAAGTTAAAGA 2186
 QY 121 AATTTGAGATCGAATATAATGTAGAGCAGAGAGACCTTCAAAAACACTGCACCCCTGGGGCCT 180
 Db 2187 AATTCGAGATCGAATACACGTAGAGCAGAGAGCTTCGAAACACTGGACCCCTGGGGCCT 2246
 QY 181 CTTAGCCCAATGAGTGCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATT 240
 Db 2247 CTTAGCCCAATGAGTGCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATT 2306
 QY 241 TTTACTCTCTTTGAGACCTGTATCTTCAACTCTCTCTTGAAGTTTGTCTTCCAGAAAT 300
 Db 2307 GCTACTCTCTTTGAGACCTGTATCTTAAACCTCTCTTGTAACTTTGTCTTCTTCCAGAAAT 2366
 QY 301 TCAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 360
 Db 2367 CGAAGCTGTAAAGCTA-----CAAAATGGAGCCCAAGATGCGAGTCCCAAGACTAA 2414
 QY 361 AATCTACCGTGGACCCCTGGAGCCGCTGCTAGACTATGCTCTGTATGTTAATGACATTGA 420
 Db 2415 GATCTACCGCAGACCCCTGGAGCCGCTGCTAGCCACGATCTGATGTTAATGACATCAA 2474
 QY 421 AGTCACCCCTCCCGAGGAATCTCAATGTCACAAACCCCTACTACACTCCCAATTCAGTAGG 480
 Db 2475 AGGCACCCCTCTCTAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 2534
 QY 481 AAGCAGTTAGACAGTTGTGACCAACCTCCCAACAGTACTTGGGTTTTCTTGTGAGA 540
 Db 2535 AAGCAGTTAGACAGGTCGTCGCGCAACCTCCCAACAGCAGCTTAGGTTTTCTTGTGAGA 2594
 QY 541 GGTGGAGTGTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATTCNNAAAGCCT 600
 Db 2595 TGGGGAGTGTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATTCNNAAAGCCT 2654
 QY 601 ANCTGGGAAGTGTGACCGCATCCATCTTTAAACATGCGGCTTGCACCTTAGCTCACACCCG 660
 Db 2655 AGCTGGGAAGTGTGACCATCCATCTTTAAACATGCGGCTTGCACCTTAGCTCACACCTG 2714
 QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
 Db 2715 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 2774
 QY 721 AATCATCTAATTCCTGTAGAGCAAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780

Db	2775	AATCATCTATTGCTGAGAGCAGCAGAGGAGCAATGATCGGGATATAAAACCAAGTC	2834
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	840
Db	2835	TTGAGCCGGCAACGGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	2894
Qy	841	CACCTATTTCACCTCTATTAAATCATGCAACTGCA	875
Db	2895	CATGCTATTTCACCTCTATTAAATCTTGCAACTGCA	2929
RESULT 4			
US-09-949-016-5640			
; Sequence 5640, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5640			
; LENGTH: 2763			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-5640			
Query Match			
Best Local Similarity 54.9%; Score 729.2; DB 4; Length 2763;			
Matches 792; Conservative 0; Mismatches 70; Indels 12; Gaps 1;			
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGCTAACCCGCCAAAAGAGGGGAACTGTTATTTTT	60
Db	1902	TCAAAATCGAAGAGCTTTAGACTTGCTAACCCGCCAAAAGAGGGGAACTGTTATTTTT	1961
Qy	61	AGGGGAAGATCTGTTAGTATGTTAACTGAATCTGGAATCATCTAGTGAAGTTAAGA	120
Db	1962	AGGGGAAGATCTGTTATTTATTTAATCAATCCGGATCGTCACTGAGAAGTTAAGA	2021
Qy	121	AATTTGAGATCGAATATAATGATAGAGCAGAGGACCTTCAAAACACTGCAACCTGGGGCT	180
Db	2022	AATTCGAGATCGAATACACGATAGAGCAGAGGAGCTTCGAACACTGGACCTGGGGCT	2081
Qy	181	CCTCAGCCAAATGGATGCCCTGAGACTCTCCCTTTCTAGGACCTCTAGCAGCTATAATTT	240
Db	2082	CCTCAGCCAAATGGATGCCCTGAGACTCTCCCTTTCTAGGACCTCTAGCAGCTATAATTT	2141
Qy	241	TTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTGTTAAGTTTGTCTCTCCAGAT	300
Db	2142	GCTACTCTCTTTGGACCCCTGATCTTAACTCTCTGTTAAGTTTGTCTCTCCAGAT	2201
Qy	301	TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCACTCACTAA	360
Db	2202	CGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCACTCACTAA	2249
Qy	361	AATCTACCGTGGACCCCTGGACCGGGCTGCTAGACTGATGCTCTGATGTTAATGACATTGA	420
Db	2250	GATCTACCGCAGACCCCTGGACCGGGCTGCTAGCCACGATCTGATGTTAATGACATCAA	2309
Qy	421	AGTCACCCCTCCCGAGGAAATCTCAACTGCACAACCCCTACTACACTCCAAATTCAGTAGG	480
Db	2310	AGGCACCCCTCCCGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG	2369
Qy	481	AAGCAGTTAGCAGGTTGTGACGCCAACCTCCGCCAACAGTACTTTGGGTTTTCTGTTGAGA	540
Db	2370	AAGCAGTTAGGCGTCTGCGCAACCTCCCAACAGCAGCTTAGGTTTTCTGTTGAGA	2429
Qy	541	GGGTGAGCTGAGAGCAGGAGCTAGCTGGATTTCTTAGCTGAGTAAAGATCCCAAGCCT	600
Db	2430	TGGGGACTGAGAGCAGGAGCTAGCTGGATTTCTTAGCTGAGTAAAGATCCCAAGCCT	2489
Qy	601	ANCTGGAGGTGACCGCATCTTAAACATGGGCTTGCAACTTAGCTCACACCCG	660
Db	2490	AGCTGGAAAGGTGACCATCTCCACCTTTAAACAGGGGCTTGCAACTTAGCTCACACCTG	2549
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCCAAAAACAGGAGGTAAAGCAATAGCC	720
Db	2550	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCCAAAAACAGGAGGTAAAGCAATAGCC	2609
Qy	721	AATCATCTATTGCTGAGAGCAGCGGAGGAGGAGGATTCGGATATATAAATCAGGCA	780
Db	2610	AATCATCTATTGCTGAGAGCAGCGGAGGAGGAGGATTCGGATATATAAATCAGGCA	2669
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	840
Db	2670	TTGAGCCGGCAACGGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	2729
Qy	841	CACCTATTTCACCTCTATTAAATCATGCAACTGC	874
Db	2730	CATGCTATTTCACCTCTATTAAATCTTGCAACTGC	2763
RESULT 5			
US-09-949-016-13002/c			
; Sequence 13002, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 13002			
; LENGTH: 168394			
; TYPE: DNA			
; ORGANISM: Human			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (1)...(168394)			
; OTHER INFORMATION: n = A,T,C or G			
US-09-949-016-13002			
Query Match			
Best Local Similarity 54.8%; Score 728.4; DB 4; Length 168394;			
Matches 878; Conservative 0; Mismatches 134; Indels 21; Gaps 5;			
Qy	257	CCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAATGAGCTGTAAGCTA	316
Db	25284	CCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAATGAGCTGTAAGCTA	25205
Qy	317	CAATAGTTCTTCAAAATGGAACCCAGATGCACTCCATGCTAAATCTACCGTGGACCC	376
Db	25204	CAATAGTTCTTCAAAATGGAACCCAGATGCACTCCATGCTAAATCTACCGAGACCC	25145
Qy	377	CTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTAAGTCACTCCCTCCGAG	436
Db	25144	--GGACCAGCCTGCTAGCCCATGCTC--CAAGTTAATACATCAAGGACCTCTCCAGAG	25088


```

RESULT 7
US-09-949-016-12249/c
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

```

```

RESULT 8
US-09-949-016-15858/c
; Sequence 15858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15858
; LENGTH: 145320
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15858

Query Match          48.4%;   Score 643.4;   DB 4;   Length 145320;
Best Local Similarity 85.4%;   Pred. No. 3.2e-208;

```

	Matches	781;	Conservative	0;	Mismatches	88;	Indels	46;	Gaps	4;
Qy	1	TCAAAATCGAAGAGCTTTAGACTTCTGCTAACCGCCAAAAAGAGGGGAACCTGTTTATTTTT	60							
Db	92888	TCAAAATCGAAGAGCTTTAGACTTCTGCTAACCTCTGAAAGAGGGGAAGCTGTTTATTTTT	92829							
Qy	61	AGGGGAAGAATGCTGTGTAG-----TATGTTAATCA	90							
Db	92828	AGGGGAAGAATGTTGTTATTTATGTTATTTTAGCGGAAGAATGTTGTTATTTATGTTAATCA	92769							
Qy	91	ATCTGGAATCATTACTTGAGAAAGTTTAAAGAAATTTTGAGATCGAATATATATCTAGACAGAGA	150							
Db	92768	ATCCTGAATTTGTCACAGAGAAAGTTTGAAGAAATTCGAGATTGAATACAACTGAGACAGAGA	92709							
Qy	151	GGACCTTC-AAAACACTGCACCTCGGGGCTCCTCAGCCAAATGGATGCCCTGGACTCTCC	209							
Db	92708	GGAGCTTCAAAAACACAGACCTCGGGGCTCCTCAGCCAAATGGATGCCCTGGATTCTCC	92649							
Qy	210	CTTCTTTAGGACCTCTAGCAGCTATAATTTTTTACTCTCTTTTGGAGCCCTGTATCTTCA	269							
Db	92648	CTTCTTTAGGATCTCTAGCAGCTTAATATGATACTCTCTCTTTGGAGCCCTGTATCTTTA	92589							
Qy	270	ACTTCTCTGTTAAGTTTGCTCTTCCAGAAATTTGAAGCTGTAAGCTACAAATAGTCTTCTC	329							
Db	92588	ACCTCTCTGTTAAGTTTGCTCTTCCAGAAATCAAAGTTGTAAGCTACAAATCGTCTTCTC	92529							
Qy	330	AAATGGAACCCAGATGCGAGTCCATGACTAAAATCTACCGTGGACCCCTGGACCGGCGCTG	389							
Db	92528	AAATGGAACCCAGATGAAGTCCATGACTAAGATCTACCGTGGACCCCTGGACCGGCGCTA	92469							
Qy	390	CTAGACTATGCTCTGATGTTTAATGACTTGAAGTCAACCCCTCCGAGGAAATCTCAACTG	449							
Db	92468	CTAGCCCATGCTCCAATTTGTTAATGATATCGAACGCAACCCCTCCGAGGAAATCTCAACTG	92409							
Qy	450	CACAAACCCCTACTPACACTCCAATTCAGTAGGAAGCAGTTAGAGCAGTTGTCTAGGCCAACCT	509							
Db	92408	CACAAACCCCTACTATGCCCAATTCGCGAGGAAGCAGTTAGACTGCTGTCAGGCCAACCT	92349							
Qy	510	CCCCAACAGTACTTTGGGTTTTCTGTTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGA	569							
Db	92348	CCCCAACAGCACTTGGGTTTTCTGTTTGAGTGGGGGACTGAGAGACAGGATTAGCTGGA	92289							
Qy	570	TTTTCTTAGGCTGACTAAGATCCNAGCCTTANCTGGGAAGGTGACCGCATCATCTTTTA	629							
Db	92288	TTTTCTTAGGCCGACTAAGATCCCAAGCCTAGCTGGGAAGGTGACCAATCCACCTTTA	92229							
Qy	630	AACATGGGGCTTGCAACTAGCTCACACCCGACCAATC-----AGAGAGCTCACCTA	680							
Db	92228	AACATGGGGCTTGCAACTAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCTCACCTA	92169							
Qy	681	AAATGCTAATCAGGCAAAAACAGAGAGGTAAAGCAATAGCCAATCATCTATTCGCTGAGAG	740							
Db	92168	AAATGCTAATAGACAAAAACAGAGAGGTAAAAAATAGCCAATCATCTATTCGCTGAGAG	92109							
Qy	741	CACAGCGGGAAGCAGAGGATTCGGATATAAATCTCAGGCAATTCAGCCAGCAACAGCAAC	800							
Db	92108	CACAGCGGGAAGCAGATGATCGGGATATAAACCCAGGCAATTCAGCCGCGCAACGGCTAC	92049							
Qy	801	CCCCCTTTGGGTCCTCCCAATGTTATGGGAGCTCTGTTTTTCACTCTATTTCACTCTATTA	860							
Db	92048	CTTCTTTGGGTCCTCCCTTTGTTGGGAGCTCTCTCT-----GTCCTTCACTCTATTA	91995							
Qy	861	AATCATGCAACTGCA	875							
Db	91994	AATATTGCAACTGCA	91980							

RESULT 9

US-09-573-080A-385

US-09-373-080A-383
: Sequence 385, Application US/09573080A

; sequence 363, Applicant No. 6828097
; Patent No. 6828097; FACILITY NO. 002005;
; GENERAL INFORMATION:

APPLICANT: JOAN, KNOLL

11250-11259

```

; APPLICANT:  ROGAN, PETER
; TITLE OF INVENTION:  SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF
; FILE REFERENCE:  30307
; CURRENT APPLICATION NUMBER:  US/09/573,080A
; CURRENT FILING DATE:  2000-05-16
; NUMBER OF SEQ ID NOS:  479
; SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 385
; LENGTH:  780
; TYPE:  DNA
; ORGANISM:  Homo sapiens
; FEATURE:
; NAME/KEY:  rerppeat_region
; LOCATION:  (1)..(780)
; OTHER INFORMATION:  ltr17
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:  Jurka, J; Walichewicz, J; Milosavljevic, A
; TITLE:  Prototypic sequences for human repetitive DNA
; JOURNAL:  Journal of Molecular Evolution
; VOLUME:  35
; ISSUE:  4
; PAGES:  286-291
; DATE:  1992-10-
; DATABASE ACCESSION NUMBER:  Database of repetitive elements (repbase)
; DATABASE ENTRY DATE:  1996-01-26
; DATABASE ENTRY DATE:  1996-01-26
; US-09-573-080A-385

```

Query Match	48.0%;	Score 637.8;	DB 4;	Length 780;
Best Local Similarity	90.0%;	Pred. No. 6.8e-208;		
Matches 711; Conservative	0;	Mismatches 60;	Indels 19;	Gaps 2;

	Qy	549	TGAGACAGGACTAGCTGGATTTCCTTAGCGTCACTAAGAATCCNAAAGCCTANTCTGGGA	608
	Dd	1	TGAGACAGGACTAGCTGGATTTCCTTAGCGCGCACTAAGAATCCCTAAGCCTTAGCTGGGA	60
	Qy	609	AGGTGACCGCATCCATCTTTTAAAACATGGGCCTTGCAACTTAGCTTCACACCCGACCAATC-	667
	Dd	61	AGGTGACCGCTTTCACCTTTTAACA CGGGCTTGCAACTTAGCTTCACACCCGACCAATCA	120
	Qy	668	----- AGAGAGCTCACATAAATGCTAATCAGGCCAAAACAGGAGGTTAAAGCAATAGC	719
	Dd	121	GGTAGTAAAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGGAGGTTAAAGAAATAGC	180
	Qy	720	CAATCATCTATTGGCTGAGAGCACAGCGGGNAAGGACAAGGATTTGGGATATAAACTCAGGC	779
	Dd	181	CAATCATCTATTGGCTGAGAGCACAGCGGGAGGACAAATGATCGGATATAAACCCGAGGC	240
	Qy	780	ATTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCACTCTATATGGGAGCTCTGTFTT	839
	Dd	241	ATTCGAGCCGGCAACGGCTACCCTTTTGGTCCCTCCCTTTGTATGGGAGCTCTGT - -	298
	Qy	840	TCACCTATTATTTCACTCTATTAAATCATGCAACTGCATCTTCTGGTCGGTTTTTTATG	899
	Dd	299	----- TTTCACTCTATTAAATCTTTGCAACTGCATCTTCTGGTCGGTTTTGTGTAGC	350
	Qy	900	GCTCAAAGCTGAGCTTTTGTTCGCCATCCA CCACTGCTGTTTGGCCACCGTCA CAGACCCGC	959
	Dd	351	GCTCGAGCTGAGCTTTTCGCTCACCGTCCACCACTGCTGTTTGGCGCGGTTCG CAGACCCGC	410
	Qy	960	TGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTTGATCCAGCGAGG	1015
	Dd	411	CGCTGACTTTCATCCCTCCGGATCCGGCAGGGTGTCCGCTGTCTCTGATCCAGCGAGG	470
	Qy	1020	TACCATTGCCACTCCCGATCAGGCTAAAGGCTTGGCAATTTGTTCTCTGCATGGCTAAGTGC	1079
	Dd	471	CGCCCATTTGCGCTCCCGATCGGGCTAAAGGCTTGGCAATTTGTTCTCTGCACGGCTAAGTGC	530
	Qy	1080	CTGGGTTTGTCTTAATAGAACTGAACACTGGTGCTACTGGTTTCCATGGTTCTCTTCCATGA	1133
	Dd	531	CTGGGTTTGTCTTAATAGAACTGAACACTGGTGCTACTGGTTTCCAGGGTCTCTCTTCCGTA	590

RESULT 9

US-09-573-080A-385

US-09-373-080A-383
: Sequence 385, Application US/09573080A

; sequence 363, Applicant No. 6828097
; Patent No. 6828097; FACILITY NO. 002005;
; GENERAL INFORMATION:

APPLICANT: JOAN, KNOLL

11250-11259


```
Qy 1140 CCCACGGCTTCTAATAGAGCTATAACACTACCGCATGGCCCAAGATTCCATTCTTGGT 1199
Db 591 CCCACGGCTTCTAATAGAGCTATAACACTACCGCATGGCCCAAGATTCCATTCTTGGG 650
Qy 1200 ATCTGTGAGGCCAAGAAACCCAGGTTCAGAGAANGTGAGGCTTGGCCACCATTTGGGAAGTG 1259
Db 651 ATCCGTGAGGCCAAGAAACCCAGGTTCAGAGAANGTGAGGCTTGGCCACCATTTGGGAAGTG 710
Qy 1260 GCCACATGCCATTTTGTAGGCGGCCACCAACATCTTGGGAGCTGTGGGAGCAAGATCC 1319
Db 711 GCCCGCGCCATTTTGGAAAGCGCGCCGACCATCTTGGGAGCTCTGGGAGCAAGAAAC 770
Qy 1320 CCCAGTAACA 1329
Db 771 CCCGTAACA 780

RESULT 10
US-09-949-016-15497/c
; Sequence 15497, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15497
; LENGTH: 245286
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(245286)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-15497

Query Match 47.8%; Score 634.6; DB 4; Length 245286;
Best Local Similarity 89.3%; Pred. No. 4.8e-205;
Matches 693; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

Qy 545 GGACTGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGATCCCAAGGCTTANCT 604
Db 142456 GTAGTGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGATCCCTAAGCCTAGCT 142397
Qy 605 GGGAGGTTGACGGCATCCATCTTTAAACATGGGCTTCCAACTTAGCTCACCACCGACCA 664
Db 142396 GGGAGGTTGACTTCATCCACCTTTTAAACATGGGCTTCCAACTTAGCTCACCACCGACCA 142337
Qy 665 ATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCCCAATC 724
Db 142336 ATCAGAGAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAGCAATAGCCCAATC 142277
Qy 725 ATCTATTGCTTGGAGACAGCGGAGGACCAAGATTGGATATAAATCAGGCAATCA 784
Db 142276 ATCTATCCCTTGGAGACAGAGGAGGACCAATGATCAGGATATAAACCAGGCAATCC 142217
Qy 785 AGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTCACT 844
Db 142216 AGCCAGCAACCGGTACACTCTTTGGGTCCCTCCCTTTATATGGAGCTCTGTTTCACT 142157
Qy 845 CTATTCTACTCTATTAATATCAATGCACTCTTCTGGTCCGTTGTTTATAGGCTCA 904
Db 142156 CTATTCTACTCTATTAATCTTGAACCTGCACTCTTCTGGTCCATGTTTGTAGGGCTCA 142097
```

```
Qy 905 AGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTCCACCGTTCACAGACCCGCTGCTG 964
Db 142096 AGCTGAAGTTTCGCTCG-CGTCCACCACTGCTGTTTCCCGTGTTCACAGACCTGCGCTG 142038
Qy 965 ACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCGAGGTACCC 1024
Db 142037 ACTTCCATCCCTCCAGATCAGGCGAGGTTGCCACTGTGCTCTCTGATCCAGCGAGGTGCC 141978
Qy 1025 ATTGCCACTCCGATCAGGCTTAAAGGCTTGGCATTTGCTGCTGATGCTAAGTGCCTGGG 1084
Db 141977 ATTGCCACTCCTGATCTGGCTTAAAGGCTTGGCATTTGCTGCTGATGCTAAGTGCCTTAGG 141918
Qy 1085 TTTGTCTCTAATAGAATGAACTGTGTCACCTGGGTTCCATGGTTCTCTTCCATCACCCAC 1144
Db 141917 TTCATCTAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCGTGACCCAC 141858
Qy 1145 GGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTATCTG 1204
Db 141857 AGCTTCTAATAGAGCTATAACACTCACCTCATGGCGCAAGATTCGTTCTTGGAAATCA 141798
Qy 1205 TGAGGCCAAGAACCCAGGTCAGAGAANGTGAGGCTTGGCCACCATTTGGGAAGTGGCCCA 1264
Db 141797 TGAGGCCAAGAACCTCAGGTCAGAGAACATAGGCTTTGCAACCATTTTGGGAAGCGGCTG 141738
Qy 1265 CTGCCATTTTGTAGCGGCCACCAACCATCTTGGGAGCTGTGGGAGCAAGGATCCC 1320
Db 141737 CCGCCATTTTGGAAAGTGGCCTGCCACCATCTTAGGAGCTCTGGGAGCAAGGACCCC 141682

RESULT 11
US-09-949-016-15086/c
; Sequence 15086, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15086
; LENGTH: 219964
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15086

Query Match 43.8%; Score 582; DB 4; Length 219964;
Best Local Similarity 87.3%; Pred. No. 5.2e-187;
Matches 687; Conservative 0; Mismatches 83; Indels 17; Gaps 4;

Qy 549 TGAGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGATCCCAAGGCTTANCTGGGA 608
Db 92554 TGAGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGATCCCAAGGCTTANCTGGGA 92495
Qy 609 AGTGACCGCATCCATCTTTAAACATGGGCTTGCACCTAGCTCACCACCGCAATCA 668
Db 92494 AGTGACCGCATCCATCTTTAAAGATGGGCTTGCACCTAGCTCACCACCGCAATCA 92435
Qy 669 GAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCAATCATCT 728
Db 92434 GAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAGCAATAGCTAATCATCT 92375
Qy 729 ATTGCTCTGAGAGCACAGCGGAGGACCAAGATTGGGATATAAATCAGGCAATCAAGCC 788
Db 92375 ATTGCTCTGAGAGCACAGCGGAGGACCAAGATTGGGATATAAATCAGGCAATCAAGCC 788
```


Db 92374 ATCCCTAAGACACAGCGGGGACAAATGATCGGGATATAAACCAGGCATTCAGCC 92315
Qy 789 AGCAACAGAACCCCTTTGGTCCCTCCCATTTGATGAGGAGCTCTGTTTCACTCTAT 848
Db 92314 GCAACGGCTACCTCTTTGGTCCCTCCCTTTGATGGAGCTCTGT-----C 92265
Qy 849 TTCACTCTATTAAATCATGCAACTGCAC--TCTTCTGGTCCGTTGTTTATGGCTCAAG 906
Db 92264 TTCACTCTATTAAATCTTGCAACTGCATCTTCTTTGGTCTGTGTTGTTAGCGCTCGAG 92205
Qy 907 CTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTCCACCGTCAAGACAGCCCGCTGTCAC 966
Db 92204 CTGAGCTTTGCTGCGCGTCCACCACTGCTGTTTGCACATGTTACAGACCTGTGGCTGAC 92145
Qy 967 TTCCATCCCTTTGGATCCAGGAGAGTCCACTGTCTCTGATCCAGGAGGTACCCAT 1026
Db 92144 TTCCA-CAGGGTGGATCCAGCAGGGTGTCTGTACTCTCTGATCCAGCTAGGCGCCAC 92086
Qy 1027 TGCCACTCCGATCAGGCTAAAGGCTTGCCATTTGCTGCTAGTGGTAAAGTCCCTGGGT 1086
Db 92085 TGCTGCTCCGATAGGCTTACAGGCTTGCCATCTGTTCTACATGCTAGTGGCGGGT 92026
Qy 1087 TGCTCTAATAGAACTGAACACT-----GCTCACTGGGTTCCATGTTCTTCCATGACCC 1142
Db 92025 CGTCTAATCAGCTGAACACTAGTCACTGCTGCTGCTTCCACAGTCTCTTCCGTAACC 91966
Qy 1143 ACGGCTTCTAATAGAGCTATAACACTACCGATGCGGATGCGCCCAAGATTCCTTCTGATC 1202
Db 91965 ACGGCTTCTAATAGAGCTATAACACTACCGATGCGGATGCGCCCAAGATTCCTTCTGATC 91906
Qy 1203 TGTGAGGCCAAGAACCCAGGTCAGAGAAGTGGCTTGCACCATTTGGGAAGTGGCC 1262
Db 91905 CTGAGGCCAAGAACCCAGGTCAGAGAAGTGGCTTGCACCATTTGGGAAGTGGCC 91846
Qy 1263 CACTGCAATTTGGTAGCGGCCACCAACCATCTTGGAGCTGTGGAGCAAGATCCGCC 1322
Db 91845 CGCCGCCATTTGGAGTGGCTGCCACCATCTTGGAGCTGTGGAGCAAGATCCGCC 91786
Qy 1323 AGTAACA 1329
Db 91785 GGTAACA 91779

RESULT 12
US-09-949-016-17420
; Sequence 17420, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17420
; LENGTH: 84571
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(84571)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17420

Query Match 42.6%; Score 565.6; DB 4; Length 84571;

Best Local Similarity 86.5%; Pred. No. 1.le-181;
Matches 688; Conservative 0; Mismatches 82; Indels 25; Gaps 5;
Qy 547 ACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAA-GAATCCCNAAAGCTTANCTG 605
Db 7322 ACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAAGGAATCCCAAGCTTGTG 7381
Qy 606 GGAAGGTGACCGCATCCATCTTTAAACATATGGGCTTGAACCTTAGCTCACACCCGACAA 665
Db 7382 GGAAGGTGACCGCTTCCACCTTTAAACATGGAATTTGCACTTAGCTCACACCCGACAA 7441
Qy 666 TCA-----GAGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGGTAAAGCAAT 716
Db 7442 TCAGATAGTAGGAGAGCTCACTAAATGCTTAATTTAGGCNAAACAGAGGTAAAGAAAT 7501
Qy 717 AGCCAATCATCTATTGCTGCTGAGAGCAGCGGAGGAGCAAGGATTTGGATATAAATCTCA 776
Db 7502 AGC-----CATCTGCTGCTGAGAGCAGAGGAGGAGCAATGATCAGGATATAAACC 7557
Qy 777 GGCATTTCAAGCAGCAACAGCAACCCCTTTGGGTCCTCCCTTCAATTTGATGGAGCTCTG 836
Db 7558 GGCATTTCAAGCAGCAACAGCAACCCCTTTGGGTCCTCCCTTCAATTTGATGGAGCTCTG 7617
Qy 837 TTTTCACTCTATTTCACCTCTATTAAATCATGCAACTGCACTTCTTGTGCTCGTGTGTTT 896
Db 7618 TTTTCACTCTGTTTCACTCTATTAAATCATGCAACTGCACTTCTTGTGCTCGTGTGTTT 7677
Qy 897 ATGGCTCAAGCTGAGCTTTTGTGCGCATCCACCACTGC-----TGTTTGGCACC 946
Db 7678 ACAGCTCAGCTGAGCTTTTGTGCGCATCCACCACTGC-----TGTTTGGCACC 7737
Qy 947 GTCAGACACCGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACTGTGCTCC 1006
Db 7738 ATCGCAGACCTGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACTGTGCTCC 7797
Qy 1007 TGATCCAGGAGGTACCATTCGCACTCCGATCAGGCTAAAGGCTTGCATTTGCTGCTG 1066
Db 7798 TGATCCAGGAGGTACCATTCGCACTCCGATCAGGCTAAAGGCTTGCATTTGCTGCTG 7857
Qy 1067 CATGCTCAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTTCCATGG 1126
Db 7858 CAGGCTCAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTTCCATGG 7917
Qy 1127 TTCTCTTCCATGACCCAGGCTTCTAATAGAGCTATAACACTCACGCAATGGCCCCAAGAT 1186
Db 7918 TTCTCTTCCATGACCCAGGCTTCTAATAGAGCTATAACACTCACGCAATGGCCCCAAGAT 7977
Qy 1187 TCCATTCCTTGGTAT-CTGTGAGGCCAAGAACCCAGGTCAGAGAANGTGAAGCTTGCCA 1245
Db 7978 TCCATTCCTTGGTAT-CTGTGAGGCCAAGAACCCAGGTCAGAGAANGTGAAGCTTGCCA 8037
Qy 1246 CCATTTGGAGTGGCCCACTGCGCTTTTGGTAGGGCCCCACCCATCTTGGGAGCTGT 1305
Db 8038 CCATTTGGAGTGGCCCACTGCGCTTTTGGTAGGGCCCCACCCATCTTGGGAGCTGT 8097
Qy 1306 GGGAGCAAGGATCCC 1320
Db 8098 GTGAGCAGGAGCCCC 8112

RESULT 13
US-09-949-016-15393
; Sequence 15393, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

Db 181 GAAATCTCACTGCAACACCCCTACTATGCCCAATTGACGGGAGAGAGTTAGAGCGT 240
QY 497 TGTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTGTAGAGGGTGGACTGAGAGAC 556
Db 241 CATGAGCAACCTCCCAACAGCACTTGGGTTTTCTGTGTAGAGGGGGAGTGTAGAGAC 300
QY 557 AGGACTAGTGTGATTTCTTAGCTGACTAAGATCCCAAGCCTTANCTGGGAAGGTGACC 616
Db 301 AGGACTAGTGTGATTTCTTAGCTGACTAAGATCCCAAGCCTTANCTGGGAAGGTGACT 360
QY 617 GCATCATCTTTAAACATGGGCTTGCACACTTTAGCTTCACACCCGACCAATCAGAGAGTCT 676
Db 361 GCATCATCTTTAAACATGGGCTTGCACACTTTAGCTTCACACCCGACCAATCAGAGAGTCT 420
QY 677 ACTAAATCTTAATCAGGCAAAACAGGAGTAAAGCAATAGCAATAGCAATCATCTATTGCTGT 736
Db 421 ACTAAATCTTAATCAGGCAAAACAGGAGTAAAGCAATAGCAATAGCAATCATCTATTGCTGT 480
QY 737 AGAGCACAGGGGAGGAGCAAGGATGGGATATTAACCTCAGGCATTTCAAGCCAGCAACAG 796
Db 481 AGAGCACAGGGGAGGAGCAAGGATGGGATATTAACCTCAGGCATTTCAAGCCAGCAACAG 540
QY 797 CAACCCCTTTGGGCTCCCTCCCAATGTATGGAGCTCTGTTTTCACTCTATTTCACCTCT 856
Db 541 CAACCCCTTTGGGCTCCCTCCCTTTGTATGGGCTCTGTTTTCACTCTATTTCACCTCT 600
QY 857 ATTAATCATGCAACTGCA 875
Db 601 ATTAATCATGCAACTGAA 619

RESULT 15
US-09-949-016-17002
; Sequence 17002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17002
; LENGTH: 140224
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(140224)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17002

Query Match 37.6%; Score 500.2; DB 4; Length 140224;
Best Local Similarity 82.4%; Pred. No. 4.6e-159;
Matches 631; Conservative 0; Mismatches 106; Indels 29; Gaps 4;
QY 535 TTGAGAGGGTGGACTGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCN 594
Db 16074 TAGCAAGCCCTATTTCAGACAGGAGTACTGATTTCTTAGGCCCACTAAGATCCCT 16133
QY 595 AAGCCTANCTGGGAAGGTGACCGCATCTTTTAAACATGGGCTTGCACCTTAGTCTCA 654
Db 16134 AAGCCTAGCTGGGAAGGTGACCGCATCTTTTAAACATGGGCTTGCACCTTAGTCTCA 16193
QY 655 CACCCGACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCCAAACAGGA 705

Db 16194 CACCGCAATCAGGTAGTAAGAGAGCTCACTAAATGCTAATTAGCRAAAACAGGA 16253
QY 706 GGTAAAGCAATAGCAATCATCTATTGCTGAGAGCACAGCGGGAGGACGAAGGATTGGG 765
Db 16254 GGGAAAGAAATAGCAATCATCTATTGCTGAGAGCATAGGGGGAGGACGAATTTGTCAGG 16313
QY 766 ATATAAATCTCAGGCAATCAAGCCAGCAACACCCCTTTGGGTCCTCCCTCCCAATTGTA 825
Db 16314 ATATAAATCTCAGGCAATCAAGCCAGCAACACCCCTTTGGGTCCTCCCTCCCTTTGTA 16373
QY 826 TGGGAGCTCTGTTTTCATCTATTTCATCTATTAAATCATGCAACTGCACTCTCTCTGT 885
Db 16374 TGGGAGCTCCGTTTTCATCTATTAAACCTTCCAACTGCA-----CACTCTCTCTGT 16425
QY 886 CCGTGTGTTTTATGCTCAAGCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGGCCAC 945
Db 16426 CCGTGTGTTTTATGCTCAAGCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGGCCAC 16485
QY 946 CQTCAAGACCCGCT-----GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGT 994
Db 16486 CATCCGACACCCGCGCTGACCCGCGCTGACTTCCACCTTCCAGATCCAGCAGGCTGT 16545
QY 995 COACTGTGCTCTGATTCAGGAGGTACCCATTTGCCACTCCCGATCAGGCTAAGGCTTG 1054
Db 16546 CTGCTGGGCTCTGATTCAGGAGGTGCGCATTTGCGGCTCCCGATTGGGCTTAAAGGCTGG 16605
QY 1055 CCATGTTTCTGCTGATTCAGGAGGTGCGCATTTGCGGCTCCCGATTGGGCTTAAAGGCTGG 1114
Db 16606 CCATGTTTCTGCTGATTCAGGAGGTGCGCATTTGCGGCTCCCGATTGGGCTTAAAGGCTGG 16665
QY 1115 TGGGTTCCATGTTCTCTTCCATGACCCAGCGCTTCTAATAGAGCTATTAACACTCACCG 1174
Db 16666 TGGGTTCCATGTTCTCTTCCATGACCCAGCGCTTCTAATAGAGCTATTAACACTCACCG 16725
QY 1175 ATGGCCCAAGATTCCATTCCTTTGGTATCTGTGAGGCAAGAACCCAGCTCAGAGAGT 1234
Db 16726 GTGGCTTAAGATTCCATTCCTTTGGAATCCCTGAGGCCAAGAA-CCAGGTCAGAAAAACAA 16784
QY 1235 GAGGCTTGCCCAATTTGGGAGTGGCCCACTGCGCAATTTGGTAGC 1280
Db 16785 GAGGCTTGCCCAATTTGGGAGTGGCCCACTGCGCAATTTGGTAGC 16830

RESULT 16
US-09-949-016-13413/c
; Sequence 13413, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13413
; LENGTH: 116652
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(116652)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13413

```
Query Match      36.0%; Score 478.6; DB 4; Length 116652;
Best Local Similarity 82.2%; Pred. No. 1.1e-151;
Matches 671; Conservative 0; Mismatches 102; Indels 43; Gaps 9;

QY 547 ACTGAGACAGGACTAGCTGATTTCTAGGCTGACTAAGAATCCCNAAAGCTTANCT-G 605
Db 67040 AGTGAGACAGGACTAGCTGATTTCTAGGCTCACTAAGAAATTCCTAAGCCTAGCTGG 66981

QY 606 GGAAGTGACCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAA 665
Db 66980 GGAAGTGACCGCACCCACCTTTAAACAAGGGGCTTGAACGCGAGTTACACACCTGACCAA 66921

QY 666 TC-----AGAGAGCTCACTAAATGCTAATCAGCG-AAAAACAGGAGGTAAGCAA 715
Db 66920 TCAGGTAGTAAGAGGGCTCACTAAATACCAATTAGGCTAAAGCAGGAGGTAAGAAA 66861

QY 716 TAG-CCAATCATCTATTTCCTGAGAGCACAGCGGGAAGGACAAGATTGGGATATAAA-C 773
Db 66860 TAGTCAAAATAATCTATCATCTGAGAGCACAGGGGAGGGACAATGATTGGGATATAAAC 66801

QY 774 TCAGGCATTCAAGCCAGCA-ACAGCAACCCCTTTGGGTCCCTCCCAATTGTATGGGAGC 832
Db 66800 CAGGCATTGCGAGTGGGAGTGGGCAACCCCTTTGGGTCCCTCCCAATTGTATGGGAGC 66741

QY 833 TCTGTTTTCACTCTATTTCACTCTATTAAATCATGAACTG--CACTCTTCTGGTCCGTG 890
Db 66740 TCTGT-----TTTCACTCTATTAAATCTTGCAACTGCACAGGCTTCTGGTCCGTG 66691

QY 891 TTTTATTTGGGTCAAGCTGAGCTTTTGTTCGCAATCCAACCTGCTGTGTTGCAACCGTCA 950
Db 66690 TTTGTTAGGGTTCGAGCTGAGCTTTCACTCGCTGTGCCACCACTGCTGTTGCCACCATCC 66631

QY 951 CAGACCCCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACCTGTGCTCCCTGAT 1010
Db 66630 CAGACCCCACTTCTGACTTCCACCCCTCCGGATCCGGCAGGGGTGCCACTGTGCTCCTGAT 66571

QY 1011 CCAGCGAGGTACCCATGTCACCTCCCGATCAGGCTTAAAGGCTTGCCATTGTTCCCTGATG 1070
Db 66570 CCAGCGAGCACCATTCGCTCCCGATCAGGCTAGAGGCTCGCCATTGTTCCCTGATG 66511

QY 1071 GCTAAGTGCCTGGGTTGCTCTAATAGAACTGAACACTGGTCACTGGGTTTCATGGTTCT 1130
Db 66510 GCTAAGTGCCTGGGTTCTATCTCTAATCAAGCTGAACACTAGTGCCTGGGTTCCACGGTTCT 66451

QY 1131 CTTCCATGACCCACCGCTTCTAATAGAGCTATAACACTCACCGCATGSCCAAGATTCCA 1190
Db 66450 CTTCTGTACCCATGGCTTCTAATAGAGCTATAACACTCACCGCATGSCCAAGATTCCA 66391

QY 1191 TTCTTTGGTATCTGTGAGGCCAAGAACCCCGAGGTGAGAGAAAGTGAGGCTTGCCACCAT 1250
Db 66390 TTCTTTGGAATCCATGAGGCCAAGAACCCCGAGGTGAGAGAAAGGCTTGCTGCCATC 66331

QY 1251 TGGGAAGTGGCC-----CACTGCCATTTTGGTAGCGGCCCAACCAT 1293
Db 66330 TTGGGAGTGGCGGCCCATTTTGGGAGTGGCGGCCCATCTTTGGGAGAGCCTTCCACCAT 66271

QY 1294 CTTGGGAGCTGGGAGCAAGGATCCCCAGTAACA 1329
Db 66270 CTTGGGAGCTCTAAGAGCAAGATCCCTCTGGTAACA 66235
```

RESULT 17
US-09-949-016-11757/c
; Sequence 11757, Application US/09949016
; Patent No. 681339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11757  
; LENGTH: 131631  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(131631)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-11757
```

```
Query Match      35.3%; Score 468.6; DB 4; Length 131631;  
Best Local Similarity 83.3%; Pred. No. 3.1e-148;  
Matches 645; Conservative 0; Mismatches 102; Indels 27; Gaps 9;  
QY 545 GGACTGAGACAGGACTAGCTGATTTCTTAGGCTGACTAAGAATCCCNAAAGCTTANCT 604  
Db 29989 GAAGTGAGACAGAACTAGATGGATTTCTTAGGCCGACTAAGAATTTCTTAAGCCTAGCT 29930  
QY 605 -GGGAAGGTGACCCGATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACC 663  
Db 29929 GGGGAAGGTGACCCGATCCATCTTTAAACACAGGGCTTGTAACTCAGCTCACACCCGACC 29870  
QY 664 AATC-----AGAGAGCTCACTAAATGCTTAATCAGGC-AAAAACAGAGGATTAAGC 713  
Db 29869 AATCAGGTAGTAAGAGGGGCTCACTAAATAACAATTTGGGCTAAAGCAGAGGTAAGA 29810  
QY 714 AATAG-CCAATCATCTATTGCCCTGAGAGCACAGCGGGAAGGACAGGATTCGGGATATAAA 772  
Db 29809 AATAGTCAAAATCATATATTGCCCTGAGAGCACAGAGGAGGGACAATGATCAGGCTATAA 29750  
QY 773 -CTCAGGCATTCAAGCCA-GCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGTATGGGA 830  
Db 29749 CCCAGGCATTTCGAGCCAGGTGTGGCAACCCCTTTGGGTCCTCTCTCTTTGTATGGGA 29690  
QY 831 GCTCTGTTTTCACTCTATTTTCACTCTATTAATCATGCAACTG--CACTCTTCTGTTCCG 888  
Db 29689 GCCCTGT-----TTTCACTCTATTAATTTGCAACTGCACACTCTCTCTGTTGTTCCG 29640  
QY 889 TGTTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCACTCCACACTGCTGTTTGCCACCGT 948  
Db 29639 TGTTTGTTCTGGCTCGAGCTGAGCTTTTCAATTCAAGTCCACCACTGCTGAATGCCACAT 29580  
QY 949 CACAGACCCGCTGCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTG 1008  
Db 29579 CACAGACCCACCATTTGACTCCACCTCTCCGATCCAGCAGGGTGTCCACTGACCTTCTG 29520  
QY 1009 ATCCAGCGAGGTACCCATTTGCCACTCCCGATCAGGCTTAAGGCTTGCCATTTGTTCTGCA 1068  
Db 29519 ATCCAGCGAGGTGCCCATCTCGCACTCTCTGATCGGGCTAGAGGCTCGCCATTTGTTCTGCA 29460  
QY 1069 TGGCTAAGTGCCTGGGTTTGTCTTAATAGAACTCAACACTGCTCACTGGGTTCCATGTT 1128  
Db 29459 TGGCTAAGTGCCTAGGTTTCAAGCTTCAAGCTTCAACTAGTTGTTGGGTTCCACGGTT 29400  
QY 1129 CTCTTCCATGACCCGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCCAAGATTC 1188  
Db 29399 CTCTTCCATGACCCGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCCAAGGTTTC 29340  
QY 1189 CATTCCTTGGTATCTGTGAGCCCAAGAACCCCGAGTCAGAGAANGTGGGCTTGCCACCA 1248  
Db 29339 CATTCCTTGGAACTGTGAGCCCAAGAACCCCGAGTCAGAGAACAAGAGGCTTGCCACCA 29280  
QY 1249 T-TTGGGAAGTGGGCCCACTGCAATTTTGGTAGCGGCCCAACCACTCTCTTGGGAG 1301  
Db 29279 TCTTGGGAGTGGGCCCAACCACTCTTGGGAGTGGGCCCAACCACTCTCTTGGGAG 29226
```

RESULT 18
 US-09-949-016-17068/c
 ; Sequence 17068, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17068
 ; LENGTH: 89584
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-17068

 Query Match 35.2%; Score 467.2; DB 4; Length 89584;
 Best Local Similarity 81.2%; Pred. No. 7.2e-148;
 Matches 639; Conservative 0; Mismatches 86; Indels 62; Gaps 6;

 QY 545 GCACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCTTANCT 604
 DB 66010 GAATTAAGACAGGAGTACTAGTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCTTANCT 65951

 QY 605 GCGAAGTCAACGCTATCTTTAAACATGCGGCTTGCAACTTAGCTCACACCCGACCA 664
 DB 65950 GCGAAGTCAACGCTATCTTTAAACATGCGGCTTGCAACTTAGCTCACACCCGACCA 65891

 QY 665 ATC-----AGAGAGTCTACTAAATGCTAATCAGCGGAGGACAGGATTTGGGATATAAATC 715
 DB 65890 ATCAGGTAGTAAAGAGAGTCTACTAAATGCTAATCAGCGGAGGACAGGATTTGGGATATAAATC 65833

 QY 716 TAGCCAAATCATCTATTGCTGAGACACAGCGGAGGACAGGATTTGGGATATAAATC 775
 DB 65832 TAGCCAAATCATCTATTGCTGAGACACAGCGGAGGACAGGATTTGGGATATAAATC 65773

 QY 776 AGCATTCAAGCAGCAACAGCAACCCCTTTGGGCTCCCTCCCATTTGATGGAGCTCT 835
 DB 65772 AGCATTCAAGCAGCAACAGCAACCCCTTTGGGCTCCCTCCCATTTGATGGAGCTCT 65713

 QY 836 GTTTTCACTCTATTTCACTCTATTTAAATCATGCAACTGCA--CTCTTCTGGTCCGTTT 893
 DB 65712 GT-----TTTCACTCTATTTAAATCATGCAACTGCACTCTCTTCTGGTCCGTTT 65663

 QY 894 TTTATGGCTCAAGCTAGCTTTTGTGCGCATCCACACTGCTGTTTGGCCAGCTCACAG 953
 DB 65662 GTT-----GCGGTCCACCACTGCTGTTTGGCCAGCTCACAG 65627

 QY 954 ACCCGTGTGCTGCTTCCATCCCTTTGGATCCAGAGAGTGTCCACTGTGCTGATCCA 1013
 DB 65626 ACTTACTGTGCTGCTTCCATCCCTTTGGATCCAGAGAGTGTCCACTGTGCTGATCCA 65567

 QY 1014 GCGAGGTACCAATGGCACTCCCGATCAGGCTAAAGGCTTGCATTTCTCTCATGGCT 1073
 DB 65566 GCAAGGCTCCCAATGGCTGCTCCCTGATCGGG-----TATTCCTGCACAGCT 65522

 QY 1074 AAGTGGCTGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGCTCTCT 1133
 DB 65521 GAGTGGCTGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGCTCTCT 65462

 QY 1134 CCATGACCCAGGCTCTTAATAGAGCTATAAAGCTCACCGCATGGCCAGATTCATTC 1193
 DB 65461 CCATGACCCAGGCTCTTAATAGAGCTATAAAGCTCACCGCATGGCCAGATTCATTC 65402

QY 1194 CTTGTATCTGTGAGGCCAAGAACCCAGGTGAGAGAAAGTGTAGCTTTGCCACCAATTTGG 1253
 DB 65401 CTTGGAATCCGTGAGGCCAAGAACCCAGGTGAGAGAAAGTGTAGCTTTGCCACCAATTTGG 65342

 QY 1254 GAAGTGGCCCACTGCGCATTTTGTAGCGGCCCAACCACTATCTTGGAGCTGTGGAGCAA 1313
 DB 65341 GAAGCAGCTGCCACCATCTTTGGAAGCAGCCTGCCACCATCTTGGAGCTGTGGAGCAA 65282

 QY 1314 GGAATCCC 1320
 DB 65281 GGAATCC 65275

 RESULT 19
 US-09-949-016-17057
 ; Sequence 17057, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17057
 ; LENGTH: 154023
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-17057

Query Match 33.6%; Score 446; DB 4; Length 154023;
 Best Local Similarity 81.3%; Pred. No. 2e-140;
 Matches 607; Conservative 0; Mismatches 118; Indels 22; Gaps 7;

 QY 549 TCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCTTANCT-GGG 607
 DB 92234 TCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATTTCTTAACTTCTATCTGCGG 92293

 QY 608 AAGGTGACCGCATCTCATCTTTAAACATGGGCTTGAATAGCTCACACCCGACCAATC 667
 DB 92294 AAGGTGACCGCATCTCATCTTTAAACATGGGCTTGAATAGCTCACACCCGACCAATC 92353

 QY 668 -----AGAGAGTCTACTAAATGCTAATCAGGC-AAAAACAGGAGGTAAAGCAATA 717
 DB 92354 AGGTAATAAGAGAGAGTCTACTAAATGCTAATCAGGC-AAAAACAGGAGGTAAAGCAATA 92413

 QY 718 GCGAA-TCACTATTGCTGAGAGACACA-GCGGAGGAGCAAGGATTTGGGATATAAATC 775
 DB 92414 GTCAGTCTATCTGCTGAGAGTACCTGCGGAGGAGCAATGATCGGATATAAATC 92473

 QY 776 AGGCATTCAAGCAGCAACAGCAACCCCTTTGGGCTCCCTCCCATTTGATGGAGCTCT 835
 DB 92474 AGGCATTCAAGCAGCAACAGCAACCCCTTTGGGCTCCCTCCCATTTGATGGAGCTCT 92533

 QY 836 GTTTTCACTCTATTTCACTCTATTTAAATCATGCACTGCACT--TTCTGGTCCGTTT 893
 DB 92534 GTTTTCACTCTGT-----TAAATAATCTGCACTGCACTTTCTGGTCTGTGTTT 92586

 QY 894 TTTATGGCTCAAGCTGAGCTTTTGTTCGCATCCACACTGCTGTTTGGCCAGCTCACAG 953
 DB 92587 GTTCTGGCTCAAGCTGAGGTTTCCCTCGCTTCCACACTGCTGATGATCGGCATTCGAG 92646

 QY 954 ACCCGTGTGCTGCTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCA 1013

Db 92647 ACCCACCATTGACTTCCACCCCTCCAGATCGGAGGGTATCCACTGCACCTTCTGATCCA 92706
Qy 1014 GCGAGGTACCAATGGCCACTCCGATCAGGCTAAGGCTTGGCCATGTTCTCTGATGGCT 1073
Db 92707 GGGAGGCGCCCATTTGCCGCTCCGATCAGGCTAGAGGCTTCTACTGTCTCTGGTGCT 92766
Qy 1074 AAGTGCCCTGGGTTGTCTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTT 1133
Db 92767 AAGTGCCCTGGGTTGTCTCTAATAGAGCTGAAGAACTAGTGGTGGGTTCCAGGTTCTCTT 92826
Qy 1134 CCATGACCCCGGCTTCTAATAGAGCTATAACACTCACCGATGGCCCAAGATTCATTC 1193
Db 92827 CTGTGACCCATGGCTTCTAATAGACTTATAACACTCACTGATGCCCAAGTTCCATTC 92886
Qy 1194 CTTGGTATCTGTAGGCCAAGAACCCAGGTCAGAGAANGTAGGCTTGGCCACCATTTGG 1253
Db 92887 CTTGAAACTGTGAAGCCAAAGAACCCAGGTCAGAGAATGGAAGGCTTGTGTCATCTTT 92946
Qy 1254 GAAGTGCCCACTGCCATTTTGGTAGC 1280
Db 92947 GGAGAGCTGCCACCATCTTGGCAGC 92973

RESULT 20

US-09-949-016-17296
; Sequence 17296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17296
; LENGTH: 251672
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(251672)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17296

Query Match 33.5%; Score 445.2; DB 4; Length 251672;
Best Local Similarity 82.4%; Pred. No. 5.3e-140;
Matches 599; Conservative 0; Mismatches 106; Indels 22; Gaps 7;
Qy 548 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCCCAAGGCTTANCT-GG 606
Db 42623 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCTTAAAGCTAGCTGGG 42682
Qy 607 GAAGTGACCGCATCCATCTTTAAACATGGGGCTTGAACCTTAGCTCACACCCGACCAAT 666
Db 42683 GAAGTGACCGCACCCACTTTTAAACACGAGCTTGTAACTCAGCTCACACCCCAACAT 42742
Qy 667 C-----AGAGAGCTCCTAAATGCTTAATCAGGC-AAAAACAGAGGTAAGCAAT 716
Db 42743 CAGGTAGTAAAGAGGCGCTCACTAAATACAAATTAGGCTAAGAGCAGGAGTAAGAAAT 42802
Qy 717 AG-CCAATCATCTATTGGCTCAGAGCAGCGGAGGACAGAGGATTGGGATATAAA-CT 774
Db 42803 AGTAAATCATCTATCATCTAGAGCAGCGGGGAGGAGCAATGATTGGGATATAAAACC 42862
Qy 775 CAGGCATTTCAAGCCAGCA-ACAGCAACCCCTTTGGGTCCCTCCCTCCCATTTGTATGGAGCT 833

Db 42863 CAGGCATTCAGCGGAGTGGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCT 42922
Qy 834 CTGTTTTCACCTATTTTCACTCTATTAAATCATGCAACTGCACTCTTCTGCTCGTGTCTT 893
Db 42923 CTGTTTTCACCTATTTAAAGCTTGCACACTGCA-----CACTCTTCTGCTCTGTGTCTT 42974
Qy 894 TTTATGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAAG 953
Db 42975 GTTCTGCTCAAGCTGAGCTTTTCACTCACCATCCACCACTGCTGAATGCCCCATCAAG 43034
Qy 954 ACCGCTGCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCAATGCTGCTCTGATCCA 1013
Db 43035 ACCGCTGCTGACTTCTTACCCCTCCAGATCTGGCAGGCTGTGGCTCTCTGATCCA 43094
Qy 1014 GCGAGGTACCCATGGCCACTCCCGATCAGGCTAAGGCTTGGCCATTTGCTCTGATGGCT 1073
Db 43095 GCGAGGATCCGTTGCTGCTCTGATTTGGGCTAAAGGCTCACCATTGTTCTGTCACAGCT 43154
Qy 1074 AAGTGCTGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTT 1133
Db 43155 AAGTGCTGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCAGGTTCTCTT 43214
Qy 1134 CCATGACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTC 1193
Db 43215 CTGTGATCCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCTCAAGGTTCCATTC 43274
Qy 1194 CTTGGTATCTGTAGGCCAAGAACCCAGGTCAGAGAANGTAGGCTTGGCCACCATTTGG 1253
Db 43275 CTTGGAATCTGTAGGCCAAGAACCCCGGTCAGAGAAACAAAGGCCACCTCATCTTGG 43334
Qy 1254 GAAGTGG 1260
Db 43335 GAGCTGG 43341

RESULT 21

US-09-949-016-11973
; Sequence 11973, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11973
; LENGTH: 251682
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(251682)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11973

Query Match 33.5%; Score 445.2; DB 4; Length 251682;
Best Local Similarity 82.4%; Pred. No. 5.3e-140;
Matches 599; Conservative 0; Mismatches 106; Indels 22; Gaps 7;
Qy 548 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCCCAAGGCTTANCT-GG 606
Db 42633 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCTTAAAGCTAGCTGGG 42692

607 GAAGGTGACCGCATCTCTTTAAACATGCGGCTTGAACCTAGCTACACCGACCAAT 666
42693 GAAGGTGACCGCACCCACCTTTAAACACGGAGCTTGTAACTCAGCTCACCCACCAAT 42752
667 C-----AGAGAGCTCCTAAATGCTATCAGGC-AAAAACAGAGGTTAAAGCAAT 716
42753 CAGGTAGTAAAGAGGCTCCTAAATATCAAAATAGGCTAAGACAGAGGTTAAAGAAAT 42812
717 AG-CCAATCATCTATGCTGAGACGACGCGGAGGACAGGATGGGATATAAA-CT 774
42813 AGTCAATCATCTATCATCTGAGACGACGAGGAGGAGCAATGATGGGATATAAACCC 42872
775 CAGGCATTCAGCCAGCA-ACAGCAACCCCTTTGGGTCCTCCCTCCCAATGATGGAGCT 833
42873 CAGGCATTCAGCCGAGGAGTGGGCAACCCCTTTGGGTCCTCCCTCCCAATGATGGAGCT 42932
834 CTGTTTCACTATTTCACTCTATTAATCATGCAACTGCACTCTTTGGTCCGTGTTT 893
42933 CTGTTTCACTCTATTAAGCTTGCAACTGCA-----CACTCTTCTGCTCTGTTT 42984
894 TTTATGGCTCAAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTTGCACCGTCACAG 953
42985 GTTCTGGCTCAAGCTGAGCTTTTCACTCACCACTGCTGCTGAATGCCCGCATCACAG 43044
954 ACCCGTCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACCTGTGCTCTGATCCA 1013
43045 ACCCGTCTGACTTCTACCCCTCCAGATCTGGCAGGTTGGGCTGCTCTGATCCA 43104
1014 GCGAGTACCCATTCGCACTCCGATCAGGCTAAAGGCTTGCATGTTCTTGCATGGCT 1073
43105 GCGAGCATCCGTTGCTGCTCTGATTTGGCTAAAGGCTCACCATTGTTCTGCAAGCT 43164
1074 AAGTGCTGGTGTGCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGTTCTCTT 1133
43165 AAGTGCTGGTGTGCTTAATAGAACTGAACACTAGTGCCTGGGTTCCACGGTCTCTT 43224
1134 CCATGACCCACGCTTCTAATAGAGCTATAAAGCTCAGGCTGCGGCTGCGGCTGCGGCT 1193
43225 CTGTATCAGAGCTTCTAATAGAGCTATAAAGCTCAGGCTGCGGCTGCGGCTGCGGCT 43284
1194 CTGTATCAGAGCTTCTAATAGAGCTATAAAGCTCAGGCTGCGGCTGCGGCTGCGGCT 1253
43285 CTGTGAATCTGTGAGGCGCAAGAACCCCGGCTCAGAGAAACAAAGGCGCCACCTCATCTGG 43344
1254 GAAGTGG 1260
43345 GAGCTGG 43351

RESULT 22
US-09-573-080A-21
; Sequence 21, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: herv17
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A

TITLE: Prototypic sequences for human repetitive DNA
JOURNAL: Journal of Molecular Evolution
VOLUME: 35
ISSUE: 4
PAGES: 286-291
DATE: 1992-10-
DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
DATABASE ENTRY DATE: 1996-01-26
DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-21

Query Match 33.1%; Score 440.4; DB 4; Length 8523;
Best Local Similarity 89.4%; Pred. No. 2.2e-139;
Matches 490; Conservative 0; Mismatches 46; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGAGCTTCTAAACCGCCAAAGAGGGGAACTCTGTTATTTT 60
DB 7988 TCAAAATCGAAGAGCTTTAGAGCTTCTAAACCGCTAAGAGGGGAACTCTGTTATTTT 8047
QY 61 AGGGGAAGAATGCTCTAGTATGTTAATCAATCTGGAATCAATCTGAGAAAAGTTAAAGA 120
DB 8048 AGGGGAAGAATGCTCTAGTATGTTAATCAATCTGGAATCGTCACTGAGAAAAGTTAAAGA 8107
QY 121 AATTGAGATCGAATATATATGATAGAGAGAGACCTTCAAAACACTGACCCCTGGGCT 180
DB 8108 AATTGAGATCGAATATCAACCGTAGAGAGAGAGCTTCAAAACACTGACCCCTGGGCT 8167
QY 181 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 8168 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8227
QY 241 TTTACTCTCTTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 8228 GCTACTCTCTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8287
QY 301 TGAAGCTGTAAAGCTTCAAAATAGTCTTCAAAATGAAACCCAGATGCACTGCACTAA 360
DB 8288 CGAAGCTGTAAAGCTTCAAAATAGTCTTCAAAATGAAACCCAGATGCACTGCACTAA 8335
QY 361 AATCTACCGTGGACCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 8336 GATCTACCGCAGACCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8395
QY 421 AGTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTCACTGCACTGCACTGCACTGCACTG 480
DB 8396 AGGCAACCCCTCTGAGGAAATCTCAAGTGTGCACTGCACTGCACTGCACTGCACTGCACTG 8455
QY 481 AAGCAGTTAGAGCAGTTGTGAGCAGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 540
DB 8456 AAGCAGTTAGAGCAGTTGTGAGCAGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 8515
QY 541 GGTGGAC 548
DB 8516 TGGGGAC 8523

RESULT 23
US-09-949-016-17411/c
; Sequence 17411, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17411
; LENGTH: 99580
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17411

Query Match      32.8%; Score 435.8; DB 4; Length 99580;
Best Local Similarity 86.1%; Pred. No. 4.7e-137;
Matches 531; Conservative 0; Mismatches 74; Indels 12; Gaps 4;

QY 245 CTCCTCTTTGGACCCCTGATCTTCAACTTCCTTGTAAAGTTTGTCTTCCAGAAATTGAA 304
Db 6792 CTACATTTCAATCCCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAA 6733

QY 305 GCTGTAAAGCTACAATAGTTCTTCAATAGNACCCAGATGCGATCCATCACTAAATATC 364
Db 6732 GCTGTAAATATTACAAATTTGTTCTTCAATAGNACCCAGATGCGATCCATCACTAAATATC 6673

QY 365 TACCGTGGACCCCTGGACCCGCTCTAGACTATGCTCTGATGTTTAAATGACATTTGAAGTC 424
Db 6672 TACCGTGGACCCCTGGACCCGCTCTAGACTATGCTCTGATGTTTAAATGACATTTGAAGTC 6613

QY 425 ACCCTCCCGAGGAAATCTCAACTGCGACAAACCCCTACTACACTCCAATTCAGTAGGAGC 484
Db 6612 ACTCCTCTGATGAAATCTTAAACTGCGAGCCCTACTATGCCCAATACAGAGGAAAGC 6553

QY 485 AGTT-AGAGCAGTTCTGAGCCAACTCCCAACAGTACTTGGGTTTTCCTGTTGAGAGGG 543
Db 6552 ACTTAAAGAGCAGTCTGCGCCAACTCCCAACAGTACTTGGGTTTTCCTGTTGAGAGGG 6493

QY 544 TGGACTGAGAGACAGGACTAGCTGGATTTCTTAGCTGCTAAGAAATCCNNAAGCCTTANC 603
Db 6492 GGGACTGAGAGACAGGACTAGCTGGATTTCTTAGCTGCTAAGAAATCCNNAAGCCTTANC 6433

QY 604 TGGG-AGGTGACCGATCCATCTTTAAACATGGGCTTGCNACTTAGCTCACACCCGAC 662
Db 6432 TGGGAAAGGTGACCGATCCATCTTTAAACATGGGCTTGCNACTTAGCTCACACCCGAC 6373

QY 663 CAATC-----AGAGAGCTCACTTAAATGCTAATCAGGCAAAACAGGAGTTAAAGC 713
Db 6372 CAATCAGGTAGTAAAGAGGCTCACTTAAATGCTAATCAGGCAAAACAGGAGTTAAAGC 6313

QY 714 AATAG-CCAATCATCTATTGCTGAGACACAGCGGGAAGGACAGGATTGGGATATAA 772
Db 6312 AATAGTCAATCATATATCGCTGAGAGACAGGAGGGAAGGACAGGATTGGGATATAA 6253

QY 773 CTCAGGCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGC 832
Db 6252 CTCAGGCATTCAAGCCAGGAGGAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGC 6193

QY 833 TCTGTTTTCACCTATT 849
Db 6192 TCTGTTTTCACCTATT 6176

RESULT 24
US-09-949-016-202370/c
; Sequence 202370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202370
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202370

Query Match      32.7%; Score 434.8; DB 4; Length 601;
Best Local Similarity 86.0%; Pred. No. 2.9e-138;
Matches 520; Conservative 1; Mismatches 69; Indels 15; Gaps 3;

QY 690 TCAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCTATTTCCTCTGAGAGACAGCGGG 749
Db 601 TTAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCTATTTCCTCTGAGAGACAGCGGG 546

QY 750 AAGGCAAAAGGATTGGGATATAAACTCAGGCAATTAAGCCAGCAACACCCCTTTGG 809
Db 545 AGGCAAAAGGATTGGGATATAAACTCAGGCAATTAAGCCAGCAACACCCCTTTGG 486

QY 810 GTCCCTCTCCATTTGATGGAGCTCTGTTTCACTCTATTTCCTCTTATTAAATCATGCA 869
Db 485 GTCCCTCTCCATTTGATGGAGCTCTGTTTCACTCTATTTCCTCTTATTAAATCATGCA 426

QY 870 ACTGCACTCTTCTGGTCCGTTTATATGGCTCAAGCTGAGCTTTTGTTCGCCATCCAC 929
Db 425 ACTGCACTCTTCTGGTCCGTTTATATGGCTCAAGCTGAGCTTTTGTTCGCCATCCAC 366

QY 930 CACTGC-----TGTTCGCCACCGTCAAGACCCGCTGCTGACTTCATCCCTTTG 979
Db 365 CACTGC-----TGTTCGCCACCGTCAAGACCCGCTGCTGACTTCATCCCTTTG 306

QY 980 GATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCGAGGTACCCATTGCCACTCCCGAT 1039
Db 305 GATCCAGCAGAGTGTGCTGCTGCTGATCCAGCGAGGTACCCATTGCCACTCCCGAT 246

QY 1040 CAGCTAAAGCTTGGCCATTGTTCTGCTGAGTAACTGCTGGGTTTGTCTTAATAGAA 1099
Db 245 TGGGCTAAAGCTTGGCCATTGTTCTGCTGAGTAACTGCTGGGTTTGTCTTAATAGAA 186

QY 1100 CTGAACACTGCTCTCTGGTTCCATGTTCTTCCATGACCCAGCGCTTCTAATAGAGC 1159
Db 185 CTGAACACTGCTCTCTGGTTCTACGGTTCTTCTTCCATGACCCAGCGCTTCTAATAGAGC 126

QY 1160 TATAACACTCAGGAGTGGCCCAAGATTCCATTCTTTGGTAT-CTGTGAGGCCAAGAAC 1218
Db 125 TATAACACTCAGGAGTGGCCCAAGATTCCATTCTTTGGTAT-CTGTGAGGCCAAGAAC 66

QY 1219 CCAGGTGAGAGAAAGTGGGCTTCCCACTTTGGGAAGTGGCCCACTGCCATTTTGGTA 1278
Db 65 CCAGGTGAGAGAAAGTGGGCTTCCCACTTTGGGAAGTGGCCCACTGCCATTTTGGTA 6

QY 1279 GCGGC 1283
Db 5 GTGGC 1

RESULT 25
US-09-949-016-44299/c
; Sequence 44299, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```


; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44299
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44299

Query Match 32.2%; Score 427.8; DB 4; Length 601;
Best Local Similarity 83.9%; Pred. No. 7.3e-136;
Matches 511; Conservative 1; Mismatches 80; Indels 17; Gaps 2;

QY	538	AGAGGCTGACCTGAGAGACAGCAGCTAGCTGGATTTCTTAGGCTGACTAGAAATCCCNAG	597
DB	601	AGAGGGAGACTGAGAGACAGCAGCTAGCTGGATTTCTTAGGCGGACTAGAAATTCCTAAG	542
QY	598	CCTANTCTGGAAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACAC	657
DB	541	CCTAGCTGGAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCATAT	482
QY	658	CGACCAATC-----AGAGAGCTCACTAAATGCTAAATCAGGCAAAAACAGAGGT	708
DB	481	CTGACCAATCAGGTAGTAAAGAGAGTTCACTAAATGCTAACTTAGGCAAAAACAGAGGT	422
QY	709	AAAGCAATAGCCATCATCTATTGCTGAGACACAGCGGAGGAGGATATGATCAGGATA	768
DB	421	AAAGAAATAGCCATCATCTATTGCTGAGACACAGCGGAGGAGGATATGATCAGGATA	362
QY	769	TAAACTCAGGCTTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGG	828
DB	361	TAAACCCAGGCTTGAAGCTGGCAATGGCTACCTCTTTGGGTCCCTCCCTTTGATGG	302
QY	829	GAGCTCTGTTTCACTCTATTTCACCTCTATTAAATCATGCAAACTGCACTCTTCTGTCGG	888
DB	301	RAGCTCTGTTTCACTCTATTAAATCTTGAACCTGCA-----CACTCTTTGTCGG	250
QY	889	TGTTTTTATGCTCAAGCTGAGCTTTGTTGGCCATCCACCACTGCTGTTTGCCACCGT	948
DB	249	TGTTTTTATGCTGACCTGAGCTTTGCTGTTGCTATCCACCACTGCTGTTTGCCACCGT	190
QY	949	CACAGACCGCTGCTGCTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTG	1008
DB	189	TGCAGACCCACCACTGACTTCCACCTCTCTGGATCCAGCAGTGTGTCTACTGTCTCTG	130
QY	1009	ATCCAGCGAGGTACCACTTCCACTCCCGATCAGGCTAAAGGCTTCCCAATGTTCTGCA	1068
DB	129	ATCCAGCGAGGCGCCGCTGCACTCTGATCGGCTAAAGGCTTCCCAATGTTCTGCA	70
QY	1069	TGCTAAGTCCCTGGGTTGTTCTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTT	1128
DB	69	CGGCTAAGTGCCAGGTTTCTCTAATAGGAGCTGAACACTAGTCACTGGGTTCCACGGTT	10
QY	1129	CTCTTCCAT 1137	
DB	9	CTCTTCCGT 1	

RESULT 26
US-09-949-016-15019
; Sequence 15019, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15019
; LENGTH: 57507
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15019

Query Match 31.0%; Score 411.4; DB 4; Length 57507;
Best Local Similarity 79.6%; Pred. No. 7.6e-129;
Matches 592; Conservative 0; Mismatches 124; Indels 28; Gaps 8;

QY	549	TCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAGAAATCCCNAAAGCTTANCT-GGG	607
DB	44269	TCAGAGACAGGACTAGCTGGATTTCTTAGGCGGACTAGAAATGCTAGACCTAGCTGGGG	44328
QY	608	AAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATC	667
DB	44329	AAGGTGATCGCACCCACCTTTAAACACCGGGCTTGTAACTCAGCTCACACCTGACCAATC	44388
QY	668	-----AGAGAGCTCACTAAATGCTAAATGCTAATCAGG-CHAAAACAGGAGGTAAAGCAATA	717
DB	44389	AGGCAGTAAAGAGGGCTCATTTAAATAACCAATTAGGTTAAAAAGCAGGAGGTAAAGAAATA	44448
QY	718	G-CCATCATCTATTGCTGAGACACAGCGGAGGAGCAAGATTTGGGATATAAACTC-775	
DB	44449	GTCAATCATCTATCATCTGAGAGCAGGAGGAGGACATGATTTGGGATATAAACTC-44508	
QY	776	-AGGCATTTCAAGCCAGCAACCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTC-834	
DB	44509	AGGCATTTGAGCGGGAGGTGGGCCACCCCTTTGGTCCCTCCCATCTGATGGAAGCTC-44568	
QY	835	TGTTTTCACTCTATTTCACCTCTATTAAATCATGCAACTG--CACTCTCTGTCGCTGTT-892	
DB	44569	TGT-----TTTCACTCTATTAAATCTTGCACACTGCACACTCTTCTGCTCCACGTT-44618	
QY	893	TTTTTATGCTCAAGCTGAGCTTTGTTGGCCATCCACCACTGCTGTTTTCGCCACCGTCACA-952	
DB	44619	TGTTTTGCTTGAAGCTGAGCTTTGCTTCCATCTTCACTGCTGTTTTCGCCACCGTCACA-44678	
QY	953	GACCGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCC-1012	
DB	44679	GACCACTGTGACTTCCACCCCTCGGATCCAGCAGGCTGTCTGCTGCAATTTCTGATCC-44738	
QY	1013	AGCGAGGTACCAATTTGCCACTCCCGATCAGGCTAAAGGCTTGGCAATTTCTGCAATGGC-1072	
DB	44739	AGCGAGGCGCCCAATTTGCTCT--CATTCGGGCTAGAGGCTCGCCATTTGTTCTGTGCGGC-44796	
QY	1073	TAAGTGCCTGGGTTTGTCTCTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTCT-1132	
DB	44797	TAGTGTCTGAGTTTGTCTCTAATCGAGCTAAACACTAGTGTGCTGCTGCTGCTGCTGCTCT-44856	
QY	1133	TCCATGACCCACCGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTTCAAT-1192	
DB	44857	TCCATGACCCACCGGCTTCTAATAGAGCTATAACACTCACCTGATGSCCAAGGTTCAAT-44916	
QY	1193	CCTTGTATCTGTGAGGCCAAGAACCCCGAGGTGAGAGANGTGAAGCTTGCACCACTTTG-1252	
DB	44917	CCTTGGAAATCCGTGAGGCCAAGAAATGTTAGGTGAGAGAACAAAGGCTTGTGCTCCATCT-44976	
QY	1253	GGAGTGGCCCACTGCAATTTGG 1276	
DB	44977	GGAGTGGCCCACTTACCAACTTGG 45000	

RESULT 27
US-09-949-016-13633/c
; Sequence 13633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13633
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13633

Query Match      30.1%; Score 399.8; DB 4; Length 39686;
Best Local Similarity 78.1%; Pred. No. 5.7e-125;
Matches 598; Conservative 0; Mismatches 120; Indels 48; Gaps 8;

QY 549 TGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCCTANCT-GGG 607
DB 6761 TGAGAGGAGGAGGACTAGCTGGATTTCTTAGGCGGACTAGATTCCTAAGCCTAGCTGGG 6702

QY 608 AAGGTGACCGCATCCATCTTTAAACATGGGGCTTCCAACTTAGCTCAACCCGACCAATC 667
DB 6701 AAGGTGACTGCACCCACCTTTAAACACGGGGCTTGTAACTAGCTCAACCTGACCAATC 6642

QY 668 -----AGAGACTCACTAAATGCTAATCAGG-CAAAAACAGGAGGTAAAGCAATA 717
DB 6641 AGGTGGTAAAGAGGCTCACTAAATATCAATTAGGTTAAAGCAGGAGGTAAAGAAATA 6582

QY 718 -GCCAATCATCTATTGCTTGAGACACAGCGGGAAGGACAGGATTTGGGATATAAA-CTC 775
DB 6581 CTCAAATCATCTATCATCTGAGACACAGGGGGAGCGACATGATTTGGATATAAACCCC 6522

QY 776 AGGCATTTCAAGCCAGCAAGCAACCCCTTTGGGTCCCTCCCAATTGTATGGGAGCTCT 835
DB 6521 AGGCATTTGAGCAAGAGGGGCAACCCCTTTGGATCCCTACCCCTTGTATAGCAGCTCT 6462

QY 836 GTTTTCACTCTATTTCACCTCTATTAAATCATGCAACTG--CACTCTTCTGGTCCGTGTTT 893
DB 6461 GT-----TTTCACTTTTAAATCTTGAACCTGCACTCTTCTGGTCCATGTTT 6412

QY 894 TTTATGGCTCAAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTGCCACCGTCAAG 953
DB 6411 GTTCTGGCTCGAGCTGAGCTTTTGTCTCACCTTCCACCACTGCTGTTGCGCACTGTCA 6352

QY 954 ACCCGTGTGACTTCCAT-----CCCTTTGGATCCAGCAGA 990
DB 6351 ACCCGTGTTCGCCACCGTCGACACCCGCTGTGACTTCCACCCCTCGGATCTGGCAGG 6292

QY 991 GGTGTCACCTGTCTCTGATCCAGCGAGGTACCCATTGCGACTCCCGCATCAGGCTAAAGG 1050
DB 6291 GTGTCGCTGCGTTTCTGATCCAGGAGGACCCAGTGCCTTCTGTTTGGGCTAGAGG 6232

QY 1051 CTTGCCATTTGTTCTTGCACTGGCTTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGG 1110
DB 6231 CTCGCCATTTGTTTCTGTGGGCAAGTGCCAGGTTGCTCTTAATCGAGCTCAACGCTAG 6172

QY 1111 TCAGTGGGTTCCATGGTCTCTCTTCCATGACCCACCGCTCTTAATAGAGCTATAAACACTCA 1170
DB 6171 TCAGTGGGTTCCAGGATTTCTTCTTCCATGACCCATGCGCTTCTTAATAGAGCTATAACA 6112

QY 1171 CCGCATGGGCCAAGATTCCATTCTTGTGTTATCTGTGAGGCGCAAGAACCCCGAGGTCAAGA 1230
```

```
DB 6111 CTGCATGGCCCAAGGTTCCATTTCTTGGATTCATGAGCCAGAGCCTCAGGTCAAGA 6052
QY 1231 ANGTGAGGCTTGCACCATTTGGAAAGTGGCCACCTGCTGCTTTGG 1276
DB 6051 AAAAAAGGCTTGGCGCATCTTGGGAGCGCCTGCGCCCATCTTTGGG 6006

RESULT 28
US-09-949-016-11770/c
; Sequence 11770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11770
; LENGTH: 49487
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(49487)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11770

Query Match      30.1%; Score 399.8; DB 4; Length 49487;
Best Local Similarity 78.1%; Pred. No. 6.6e-125;
Matches 598; Conservative 0; Mismatches 120; Indels 48; Gaps 8;

QY 549 TGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCCTANCT-GGG 607
DB 16051 TGAGAGGAGGAGGACTAGCTGGATTTCTTAGGCGGACTAAGATTCCTAAGCCTAGCTGGG 15992

QY 608 AAGGTGACCGCATCCATCTTTAAACATGGGGCTTTCACACTTAGCTCAACCCGACCAATC 667
DB 15991 AAGGTGACTGCACCCACCTTTAAACACGGGGCTTGTAACTAGCTCAACCTGACCAATC 15932

QY 668 -----AGAGAGCTCACTAAATGCTAATCAGG-CAAAAACAGGAGGTAAAGCAATA 717
DB 15931 AGGTGGTAAAGAGGCTCACTAAATATCAATTAGGTTAAAGCAGGAGGTAAAGAAATA 15872

QY 718 -GCCAATCATCTATTGCTTGAGACACAGCGGGAAGGACAGGATTTGGGATATAAA-CTC 775
DB 15871 CTCAAATCATCTATCATCTGAGACACAGGGGGAGCGACATGATTTGGATATAAACCCC 15812

QY 776 AGGCATTTCAAGCCAGCAAGCAACCCCTTTGGGTCCCTCCCAATTGTATGGGAGCTCT 835
DB 15811 AGGCATTTGAGCAAGGAGGGGCAACCCCTTTGGATCCCTTACCTTGTATAGCAGCTCT 15752

QY 836 GTTTTCACTCTATTTCACCTCTATTAAATCATGCAACTG--CACTCTTCTGGTCCGTGTTT 893
DB 15751 GT-----TTTCACTTTTAAATCTTGGAACTGCACTCTTCTGGTCCATGTTT 15702

QY 894 TTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTGGCCACCGTCAAG 953
DB 15701 GTTCTGGCTCAGCTGAGCTTTTGTCTCACCTTCCACCACTGCTGTTGCCACTGTCAAG 15642

QY 954 ACCCGTGTGACTTCCAT-----CCCTTTGGATCCAGCAGA 990
DB 15641 ACCCGTGTTCGCCACCGTCGACAGCCGCTGTTGACTTCCACCCCTCGGATCTGGCAGG 15582
```



```

Db      234  CGTTGCTGCTCTGATTTGGGCTAAAGGCTACCAATTTGTTCTCTGCACAGCTAAGTGCTGG 175
Qy      1084  GTTTTGTCTCTAATAGAACTGAACACATGGTGCTACTGGGGTTCCATGGTTCTCTTCCATGACCCA 1143
Db      174  GTTCGTCTCTAATCAAGCTGAACACATAGTCGCTGGGTTCCACGGTTCTCTTCTGTGATCCA 115
Qy      1144  CGGCTTCTAATAGAGATTAAACATCACCAGCATGGCCCAAGATTCCAATTCCTTGGTATCT 1203
Db      114  CAGCTTCTAATAGAGCTATAACACATCACCGCATGGCTCAAGGTTCCATTCCTTTGGAATCT 55
Qy      1204  GTGAGGCCAAGAACCCCAAGGTCAGAGAANGTGAGGCTTGCCACCACATTTGGGA 1255
Db      54  GTGAGGCCAAGAACCCCGGGTCAGAGAACAAAGGCCACCTCATCTTTGGGA 3

RESULT 31
US-09-949-016-44298/c
; Sequence 44298, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCES: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; CURRENT APPLICATION NUMBER: 60/241,755
; PRIORITY FILING DATE: 2000-10-20
; CURRENT APPLICATION NUMBER: 60/237,768
; PRIORITY FILING DATE: 2000-10-03
; CURRENT APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44298
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44298

```

Query Match	27.8%	Score 369.6	DB 4	Length 601
Best Local Similarity	83.6%	Pred. No. 7e-116		
Matches 432	Conservative 1	Mismatches 76	Indels 8	Gaps 1
Qy	764	GGATATAAACTCAGGCATTCAAGCGCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTG	823	
Db	598	GGATATAAAACCCAGGCATTTGAGCTGGCAATGGCTACCCCTTTTGGGTCCCTCCCTTTG	539	
Qy	824	TATGGGAGCTCTGTTTTACCTCTATTTTCACTCTATTAATCATGCAACTGCACCTCTTCTG	883	
Db	538	TATGGGAGCTCTGTTTTACCTCTATTAATCTTCAACTGCA-----CACTCTTTTG	487	
Qy	884	GTC CGTGT TTTTATGCTCAAGCTGAGCTTTTCTGCGCAATCACCACCTGCTGTGTGGCC	943	
Db	486	GTC CGTGT TTTGTTACGGCTTGACCTGAGCTTTTCGCTTGCTATCCACACCTGCTGTGGCC	427	
Qy	944	ACCGTCAAGACCCGGCTGCTGACTTCCATCCCTTTGGATCCAGAGAGTGCACCTGTC	1003	
Db	426	ACCGTTGAGACCCACCCACCTGACTTCCACCTCTCTGGATCCAGCAGTGTGTTCACCTGTC	367	
Qy	1004	TCCTGATCCAGCGAGGTACCAATTGCCATCCCGATCAGGCTAAAGGCTTGCCCAATTGTC	1063	
Db	366	TCCTGATCCAGCGAGGCGCCGCTGCCCACTCCTGATCGGGCTAAAGGCTTGCCCAATTGTC	307	
Qy	1064	CTGCATGGCTAAAGTGCCCTGGGTTTGTCTCTAATAGAACTGAACTGGTCACTGGGTTCCA	1123	
Db	306	CTGCAYGGCTAAAGTGCCAGGTTTCATCTCTAATGGAGCTGAACCTAGTCACTGGGTTCCA	247	
Qy	1124	TGGTTCTCTCCATGACCCACGGCTTCTAATAGAGCTATAACACTCAACGATGCCCCAA	1183	
Db	246	CGGTTCTCTCCGTGACCCATGGGTTCTTAATAGAGCTATAACACTCACTGATGCCCCAA	187	
Qy	1184	GATTCCATTCTCGGTATCTGTGAGGCCAAGAACCCAGGTCAGAGAANGTGAGGCTTGC	1243	

```

Db      186  GATTCCATTCTTTGGAAACCTTTGAGGCCAAGAACCCCGAGGTCAGAGAACAGAGGCTTGC 127
Qy      1244  CACCATTGGGAAGTGGCCCACTGCCCATTTTGGTAGC 1280
Db      126  CGCTATCTTGAAGGGGCGCTGCCACCATCTTTGGGAGC 90

RESULT 32
US-09-949-016-146564
; Sequence 146564, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146564
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-146564

```

Query Match	26.0%	Score 345.4	DB 4	Length 601
Best Local Similarity	83.7%	Pred. No. 1.4e-107		
Matches 405	Conservative 1	Mismatches 68	Indels 10	Gaps 1
Qy	796	GCACCCCTTTGGGTCCCTCCCATCTGTATGGAGCTCTGTTTCACTCTATTTCACATC	855	
Db	116	GCTACCTCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTCACT-----C	165	
Qy	856	TATTAAATCATGCAACTGCACCTCTTCGTGCCGTGTTTTTATGGCTCAAGCTGAGCTTT	915	
Db	166	TATTAAATCTTGCAACTGCACCTCTTCGTGCCGTGTTTTTATGGCTCGAGCTGAGCTTT	225	
Qy	916	TGTTGGCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCTGCTGACTTCCATCCC	975	
Db	226	CACCTCCATCCACCACTGCTGTTTGGCCGCATCGCAGGCTGCGCACTGACTTCCATCCC	285	
Qy	976	TTTGGATCCAGCAGAGTGCCACTGTGCTCCTCATCCAGCGAGTACCCATTGGCACTCC	1035	
Db	286	TCTGGATCTAGCAGGKTGTCGTTGTGCTCTGTATCCAGTGAGACGCCATTGCCGATCC	345	
Qy	1036	CGATCAGGCTAAAGGCTTGGCATTTGTTCTTCGCATGGCTAAGTGCTCGGTTTGTCTTAAT	1095	
Db	346	CGACTGGCTAAAGACTTGGCATTTGTTCTTACGCGGCTAAGTGCCCGGTTTCATCTAAT	405	
Qy	1096	AGAACTGAACACTGGTCACTGGGTTTCATGGTTCTCTTCCATGACCCACGGCTTCTAATA	1155	
Db	406	TGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTTCTGTGACCCGCTGGCTTCTAATA	465	
Qy	1156	GAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCTTGGTATCTGTGAGGCCAAGA	1215	
Db	466	GAGCTATAACACTCACCGCTGGCCCAAGATTCCATTATTGGNATTCATGAGGCCAAGA	525	
Qy	1216	ACCCGAGTCAAGAAANGTGGGCTTGCCACCAATTTGGGAAGTGGCCACTGCCATTTTG	1275	
Db	526	ACCCGAGTCAAGAAACACGAGGCTTGCCATCATCTTAGAAGCAGCCGCCACCATCTTC	585	
Qy	1276	GTAG 1279		
Db	586	GGAG 589		

```

; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13590
; LENGTH: 149971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(149971)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13590

Query Match      25.9%; Score 343.6; DB 4; Length 149971;
Best Local Similarity 79.7%; Pred. No. 2.9e-105;
Matches 488; Conservative 0; Mismatches 96; Indels 28; Gaps 6;

QY 257 CCTGTATCTTCAACTTCTCTTAAAGTTTGTCTCTCCAGAAATTGAAGCTGTAAAGCTA 316
   |||||
Db 18750 CCTGTATCTTAACTCTCTTAAAGTTTGTCTCTCCAGAAATTGAAGCTGTAAAGCTA 18691

QY 317 CAAATAGTTCTTCAATGGAACCCGAGATGCGAGTCCATCAATCTACCTGGACCC 376
   |||||
Db 18690 TAAATGGTTCTTCAATGGAACCCGAGATGCGAGTCCATCAATCTACCTGGACCC 18640

QY 377 CTGGACCGGCTCTAGACTATGCTCTGATGTTTAATGACATGCAAGTCACTCCCTCCGAG 436
   |||||
Db 18639 CTAGACCGGCTCTAGCCCTAGCTCTGATGTTTAATGACATGCAAGTCACTCCCTCCGAG 18580

QY 437 GAAATCTCAACTGCACAAACCCCTACTA-CACCTCAATTTCAGTAGGAAGCAGTT-AGAGCA 494
   |||||
Db 18579 GAAATCTCAACTGCATGACCCCTACTATGCCCCCAATTTCAGTAGGAAGCAGTT-AGAGCG 18520

QY 495 GTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGAGGGTGGAGTGGAG 554
   |||||
Db 18519 GTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGAGGGTGGAGTGGAG 18460

QY 555 ACAGGACTAGCTGGATTTCTTAGGCTGACTTAAGAAATCCNAAAGCTTANCTGGGAAGTGA 614
   |||||
Db 18459 ACAAGACTAGCTGGATTTCTTAGGCTGACTTAAGAAATTCCTTAACCCAGCTGGGAAGTGA 18400

QY 615 CCGCATCCATCTTTAAACATGGGGCTTGCACACTTAGCTCACACCCGACCAATC----- 667
   |||||
Db 18399 CCGCATCCATCTTTAAACATGGGGCTTGTAACTCAGCTCACACCCGACCAATCAGATAGT 18340

QY 668 --AGAGAGTCTCACTAAATGCTTAATCAGGCAAAACAGAGGTTAAAGCAATA-----G 718
   |||||
Db 18339 AAAGAGGCTCTCACTAAATGCTTAATGCTTAAGCAAGGAGTAAAGGAGGAGGAGGAGGAG 18280

QY 719 CCAATCATCTATGCTGAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 777
   |||||
Db 18279 CAAATCATATATCCTGAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18220

QY 778 GCATTCAGCCAGCAACAGCAACCCCTTTGGGTTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 837
   |||||
Db 18219 GCATTCAGCCAGGAGTGCAACCCCTTTGGGTTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 18160

QY 838 TTTCACCTCTATT 849
   |||||
Db 18159 TCACCTCTTAAT 18148

RESULT 35
US-09-949-016-133739/c
; Sequence 133739, Application US/09949016

; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13590
; LENGTH: 149971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(149971)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13590

Query Match      25.9%; Score 344.6; DB 4; Length 443;
Best Local Similarity 88.9%; Pred. No. 2.2e-107;
Matches 394; Conservative 1; Mismatches 45; Indels 3; Gaps 2;

QY 766 ATATAAATCAGGATTCAGCCAGCAACAGCA-ACCCCTTTGGGTCCCTCCCAATGT 824
   |||||
Db 3 ATATAAACCAGGATTCAGAGCTGGCAACAGCAGCCCCCTTTGGGTCCCTCCCTTTGT 62

QY 825 ATGGGAGCTGTGTTTCACTCTATTTCACCTCTATTAAATCATGCACTGCACTCTCTGG 884
   |||||
Db 63 ATGGGAG--CTGTTTTCATGCTATTTCACCTCTATTAAATCTGCACTGCACTCTCTGG 120

QY 885 TCGGTGTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTTTTGCCA 944
   |||||
Db 121 TCCATGTTTCTACGGCTCGAGCTGAGCTTTTGTTCACCGTCCACCACTGCTTTTGCCA 180

QY 945 CCGTACAGACCCGCTGCTGACTTCATCCCTTTGATCCAGCAGAGTGTCCACTGTGCT 1004
   |||||
Db 181 CCACCGCAGACCTGCGCTGACTCCCTCTGATCTCTGATCTCTGAGGCTGCTGCTGTGCT 240

QY 1005 CCGTATCCAGGAGTACCCATTCGCCACTCCGATCAGGCTAAAGGCTTGGCAATTGTTC 1064
   |||||
Db 241 CCGATCCAGGAGGCGCCCATTTGCCGCTCCCAATTGGGCTAAAGGCTTGGCAATTGTTC 300

QY 1065 TGCATGGCTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGTTCCAT 1124
   |||||
Db 301 TGACGGCTAAGTGCCTGGGTTTGTCTTAATAGGCTGAACACTAGTCACTGGGTTCCAT 360

QY 1125 GGTTCCTTCATGACCCAGCGCTTCTTAATAGAGCTTAACACTCAGCGATGCCCCAAG 1184
   |||||
Db 361 GGTTCCTTCATGACCCAGCGCTTCTTAATAGAGCTTAACACTTACCACATGCCCCAAG 420

QY 1185 ATTCCATTCCTTGATCTGTA 1207
   |||||
Db 421 ATTCCATTCCTTGAATCCGTTGA 443

RESULT 34
US-09-949-016-13590/c
; Sequence 13590, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133739
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-133739

Query Match      25.5%; Score 338.8; DB 4; Length 601;
Best Local Similarity 89.9%; Pred. No. 2.6e-105;
Matches 372; Conservative 1; Mismatches 40; Indels 1; Gaps 1;

QY 545 GGACTGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGAAATCCNAAAGCCTANCT 604
Db 413 GTAGTGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCTAAGCCTAGCT 354

QY 605 GGGAGGTGACCGCATCCATCTTAAACATGGGGCTTCACTTGGCTCAGCCGACCA 664
Db 353 GGGAGGTGACTTCATCCACCTTTAAACATGGGGCTTCACTTGGCTCAGCCGACCA 294

QY 665 ATCAGAGCTCCTATAAATGCTATCAGGCAAAACAGGAGTAAAGCAATAGCCAAATC 724
Db 293 ATCAGAGCTCCTATAAATGCTATCAGGCAAAACAGGAGTAAAGCAATAGCCAAATC 234

QY 725 ATCTATTGCTGAGAGCAAGCGGGAAGCAAGGATTGGGATATATAAATCAGGCAATCA 784
Db 233 ATCTATCCCTGAGAGCAAGCGGGAAGCAAGGATTGGGATATATAAATCAGGCAATCC 174

QY 785 AGCCAGCAACAGCAACCCCTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 844
Db 173 AGCCAGCAACAGCAACCTCTTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 114

QY 845 CTATTTCACCTCTATTAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTATGGCTCA 904
Db 113 CTATTTCACCTCTATTAATCTTGCAACTGCACTCTTCTGGTCCATGTTTGTACGGCTCA 54

QY 905 AGCTGAGCTTTGTCGGCATCCACCACTGCTGTTTGGCAGCGTCACAGACCCG 958
Db 53 AGCTGAAGTTTCGCTCG-CGTCCACCACTGCTGTTTGGCGCTGTCCAGACCTG 1

RESULT 36
US-09-949-016-120277
; Sequence 120277, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133739
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-133739

Query Match      25.5%; Score 338.8; DB 4; Length 601;
Best Local Similarity 89.9%; Pred. No. 2.6e-105;
Matches 372; Conservative 1; Mismatches 40; Indels 1; Gaps 1;

QY 545 GGACTGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGAAATCCNAAAGCCTANCT 604
Db 413 GTAGTGAGACAGAGCTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCTAAGCCTAGCT 354

QY 605 GGGAGGTGACCGCATCCATCTTAAACATGGGGCTTCACTTGGCTCAGCCGACCA 664
Db 353 GGGAGGTGACTTCATCCACCTTTAAACATGGGGCTTCACTTGGCTCAGCCGACCA 294

QY 665 ATCAGAGCTCCTATAAATGCTATCAGGCAAAACAGGAGTAAAGCAATAGCCAAATC 724
Db 293 ATCAGAGCTCCTATAAATGCTATCAGGCAAAACAGGAGTAAAGCAATAGCCAAATC 234

QY 725 ATCTATTGCTGAGAGCAAGCGGGAAGCAAGGATTGGGATATATAAATCAGGCAATCA 784
Db 233 ATCTATCCCTGAGAGCAAGCGGGAAGCAAGGATTGGGATATATAAATCAGGCAATCC 174

QY 785 AGCCAGCAACAGCAACCCCTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 844
Db 173 AGCCAGCAACAGCAACCTCTTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 114

QY 845 CTATTTCACCTCTATTAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTATGGCTCA 904
Db 113 CTATTTCACCTCTATTAATCTTGCAACTGCACTCTTCTGGTCCATGTTTGTACGGCTCA 54

QY 905 AGCTGAGCTTTGTCGGCATCCACCACTGCTGTTTGGCAGCGTCACAGACCCG 958
Db 53 AGCTGAAGTTTCGCTCG-CGTCCACCACTGCTGTTTGGCGCTGTCCAGACCTG 1

RESULT 37
US-09-949-016-202257
; Sequence 202257, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120277
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202257

Query Match      24.4%; Score 324.6; DB 4; Length 601;
Best Local Similarity 82.0%; Pred. No. 2e-100;
Matches 387; Conservative 1; Mismatches 76; Indels 8; Gaps 1;
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120277
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-120277

Query Match      25.3%; Score 336; DB 4; Length 601;
Best Local Similarity 86.2%; Pred. No. 2.4e-104;
Matches 394; Conservative 1; Mismatches 57; Indels 5; Gaps 2;

QY 877 TCCTTCGCTCGGTGTTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCT 936
Db 16 TCCTTCGCTCGGTGTTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCT 75

QY 937 GTTTGCCACCGTCACAGACCGCTGCTGCTTCATCCCTTTGGATCCAGAGAGTGTCC 996
Db 76 GTTTGCCACCTGTTACAGACCTGCTGCTTCCA-CAGGGTGGATCCAGCAGGTGTCT 134

QY 997 ACTGTGCTCTGATCCAGCGAGGTACCCATTTGCCACTCCCGATCAGGCTAAAGCTTGCC 1056
Db 135 GCTGTACTCTGTATCCAGCTAGGCGCCCACTGCTGCTCCCGATAGGGCTACAGGCTTGCC 194

QY 1057 ATTGTTCTGCTGATGGCTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACT----GGTC 1112
Db 195 ATCGTTCTACATGGCTAAGTGCCCGGTTTCGTCCTAATCGAGCTGAACACTAGTCACTC 254

QY 1113 ACTGGGTTCCATGTTCTCTTCATGACCCAGCGCTTCTTAATAGAGCTATAAAGCTACCC 1172
Db 255 ACTGGGTTCCACAGTCTCTCTTCGTCACCCACGCGCTTCTTAATAGAGTATAAAGCTACCC 314

QY 1173 GCATGGCCCAAGATTCCATTCCTTGGTATCTGTGAGCCCAAGAACCCAGGTCAGAGAA 1232
Db 315 GCATGGCCCAAGATTCCATTCCTTGGTATCTGTGAGCCCAAGAACCCAGGTCAGAGAA 374

QY 1233 GTGAGGCTTGGCCACCATTTGGGAAAGTGGCCCACTGCCATTTTGTAGCGGCCACCA 1292
Db 375 ATGAGGCTTGGCCACCATTTGGGAAAGTGGCCCGCCCATTTTGGAAAGTGGGCTGCGACCA 434

QY 1293 TCTTGGAGCTGTGGGAGCAAGGATCCCCCAAGTAA 1329
Db 435 TCTTGGAGCTGTGGGAGCAAGGACGACCCCGTAA 471

RESULT 37
US-09-949-016-202257
; Sequence 202257, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202257
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202257

Query Match      24.4%; Score 324.6; DB 4; Length 601;
Best Local Similarity 82.0%; Pred. No. 2e-100;
Matches 387; Conservative 1; Mismatches 76; Indels 8; Gaps 1;
```



```

Db      292 CATGCCCAAGATTCCATCTTTGGAACTTTGAGGCCAAGACCCCAAGGTCAGAGACA 233
QY      1234 TGAGCTTTGCCACCAATTTGGGAAGTGGCCCACTGCCATTTTGGTAGC 1280
Db      232 AGAGCTTGGCGCTATCTTGGAAAGGGCGCTGCCACCATCTTTGGGAGC 186

RESULT 40
US-09-949-016-44296/c
; Sequence 44296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44296
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44296

Query Match      23.8%; Score 316.4; DB 4; Length 601;
Best Local Similarity 82.2%; Pred. NO. 1.3e-97;
Matches 465; Conservative 1; Mismatches 74; Indels 26; Gaps 8;

QY      424 CACCCCTCCGAGGAATCTCAACTGACACACCCCTACTACACTCCAATTCAGTAGGAAG 483
Db      600 CACTCTCTGAGGAATCTCACTGACGACCCCTACTATGCCCAATTCAGCAGGAAG 541
QY      484 CAGTT-AGAGCAGTTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGAGG 542
Db      540 CAGTTAAGAGCGTTGTCGGCCAACTCCCAACAGCAGCTTGGGTTTTCTGTTGAGATG 481
QY      543 GTGACTGAGAGACAGGACTAGCTGGATTCTTAGGCTGACTAAGAAATCCCNAAAGCCTAN 602
Db      480 GGGGACTGAGAGACAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATTCCTAAGCCTAG 421
QY      603 CT-GGGAAAGGTGACCGCATCCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCCGA 661
Db      420 CTGGGGAAGGTGACCGCACCCACCTTTAAACACAGGGTGTGTAAGTCTGAGCTCACACTGA 361
QY      662 CCAATC-----AGAGAGCTCACTAAATGCTTAATCAGGC-AAAAACAGGAGGTAAA 711
Db      360 CCAATCGGCTAGTAAAGAGGGCTCACTGAAACACAAATTAGGCTAAAGCAGGAGGTAAR 301
QY      712 GCAATAG-CCAATCATCTATTGCTTGAGAGACAGCGGGAAGGACAAGGATTGGGATATA 770
Db      300 GAAACAGTCAATCATATATATCGCTTGAGAGACACAAG-GGGACAATGATCGGGATATA 242
QY      771 AACTCAGGATTCAGACGAGCAACAGCAACCCCTTTGGGTCCTCCCTGCTATATGGA 830
Db      241 AACTCAGGATTCAGACGAGGAGTGGCAACCCCTTTGGGTCCTCCCTGCTATATGGA 182
QY      831 GCTCTGTTTTCACTCTATTTCATCTATTAAATCATGCAACTGCA--CTCTTCTGGTCCG 888
Db      181 GCTCTGT-----TTTCACTCTATTAAATCTTGCAACTGCAAGCTCTTCTGCTCTG 132
QY      889 TGTTTTTTATGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGCACCGT 948
Db      131 TGTTTTGTTACGGCTTTGAGCTGAGCTTTTGTCTCACTGTCCACACCGCTGTTTGCCTGT 72
QY      949 CACAGACCCGCTGCTGACTTTCATCC 974
```

```

Db      71 TACAGACCTGCCATTGACTCCCAACC 46

Search completed: February 21, 2005, 15:01:37
Job time : 246.166 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 11:52:50 ; Search time 734.518 Seconds
(without alignments)
10694.149 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatcgagagctttag.....gcaaggtccccagtaaca 1329

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	99.8	1329	8	US-08-979-847-108
2	1326	99.8	1329	16	US-10-114-104-108
3	1035.2	60.1	56093	10	US-09-873-367C-81
4	798.2	60.1	21646	13	US-10-087-192-910
5	778.8	58.6	1393	17	US-10-220-120-17
6	775.4	58.3	2030	17	US-10-637-565-18
7	774.2	58.3	7974	14	US-10-198-846-9936
8	766.8	57.7	2074	17	US-10-416-642-4
9	731.8	55.1	2946	13	US-10-114-893-134
10	731.8	55.1	2946	15	US-10-016-249-3
11	730.8	55.0	2930	9	US-09-902-535-1

12	722.6	54.4	7582	17	US-10-632-793-30	Sequence 30, Appl
13	721.2	54.3	1136	17	US-10-632-793-25	Sequence 25, Appl
14	712.6	53.6	2782	17	US-10-632-793-26	Sequence 26, Appl
15	711	53.5	2782	17	US-10-133-036-1	Sequence 1, Appl
16	704.8	53.0	1894	9	US-09-864-761-4444	Sequence 4444, Ap
17	643.4	48.4	161334	13	US-10-087-192-730	Sequence 730, App
18	637.8	48.0	780	10	US-09-854-867-385	Sequence 385, App
19	618	46.5	108773	19	US-10-741-600-17907	Sequence 17907, A
20	601.4	45.3	157090	18	US-10-672-764A-34	Sequence 34, Appl
21	579.8	43.6	180557	13	US-10-003-806-6	Sequence 6, Appl
22	579.8	43.6	180557	13	US-10-003-806-9	Sequence 9, Appl
23	570.2	42.9	161671	15	US-10-017-117-1	Sequence 1, Appl
24	545.4	41.0	134292	17	US-10-240-425-1102	Sequence 1102, Ap
25	545	41.0	635	8	US-08-979-847-102	Sequence 102, App
26	545	41.0	635	16	US-10-114-104-102	Sequence 102, App
27	536.2	40.3	849	17	US-10-220-120-15	Sequence 15, Appl
28	530.6	39.9	366710	18	US-10-719-993-6792	Sequence 6792, Ap
29	513.8	38.7	80032	19	US-10-741-600-17727	Sequence 17727, A
30	512	38.5	17758	17	US-10-264-237-2833	Sequence 2833, Ap
31	502	37.8	326014	9	US-09-731-231A-3	Sequence 3, Appl
32	502	37.8	326014	18	US-10-751-985-3	Sequence 3, Appl
33	478.6	36.0	283351	18	US-10-719-993-7065	Sequence 7065, Ap
34	478.6	36.0	283351	19	US-10-741-600-17995	Sequence 17995, A
35	478.6	36.0	1980090	18	US-10-719-993-6815	Sequence 6815, Ap
36	478.6	36.0	1980090	19	US-10-741-600-17676	Sequence 17676, A
37	468.8	35.3	2052	17	US-10-276-774-678	Sequence 678, App
38	465	35.0	3372	17	US-10-632-793-28	Sequence 28, Appl
39	465	35.0	22436	11	US-09-997-722-148	Sequence 148, App
40	460	34.6	285020	13	US-10-087-192-1666	Sequence 1666, Ap
41	457.8	34.4	2145	13	US-10-027-632-98772	Sequence 98772, A
42	457.8	34.4	2145	17	US-10-027-632-98772	Sequence 98772, A
43	455.8	34.3	1003	17	US-10-637-565-1	Sequence 1, Appl
44	455.8	34.3	2052	17	US-10-637-565-14	Sequence 14, Appl
45	453.6	34.1	559	9	US-09-864-761-7501	Sequence 7501, Ap

ALIGNMENTS

RESULT 1
US-08-979-847-108
; Sequence 108, Application US/08979847
; Publication No. US2003039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKER, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; AND THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435

Db	781	TTCAAAGCCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGTTTT	840
Qy	841	CACTCTATTTCACTCTATTAAATCATGCAACTGCACCTCTTCTGGTCCGTGTTTTATGG	900
Db	841	CACTCTATTTCACTCTATTAAATCATGCAACTGCACCTCTTCTGGTCCGTGTTTTATGG	900
Qy	901	CTCAAGCTGAGCTTTTGTTCGGCATCCACCACTGCTGTTTGGCACCGTCACAGACCCGCT	960
Db	901	CTCAAGCTGAGCTTTTGTTCGGCATCCACCACTGCTGTTTGGCACCGTCACAGACCCGCT	960
Qy	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGCACACTGTCTCTGATCCAGCGAGGT	1020
Db	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGCACACTGTCTCTGATCCAGCGAGGT	1020
Qy	1021	ACCCATTGGCACTCCCGATCAGGCTAAAGGCTTGCCATTGTTCTCTGCATGGCTAAGTGCC	1080
Db	1021	ACCCATTGGCACTCCCGATCAGGCTAAAGGCTTGCCATTGTTCTCTGCATGGCTAAGTGCC	1080
Qy	1081	TGGGTTTGTCTTAATAGAACTGAACTGCTGGTTCATGGGTTCCATGTTCTCTCCATGAC	1140
Db	1081	TGGGTTTGTCTTAATAGAACTGAACTGCTGGTTCATGGGTTCCATGTTCTCTCCATGAC	1140
Qy	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCTCTGGTA	1200
Db	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCTCTGGTA	1200
Qy	1201	TCTGTGAGGCCAAGAACCCAGGTCAGAGAACTGAGGCTTGCCCATTTGGGAAGTGG	1260
Db	1201	TCTGTGAGGCCAAGAACCCAGGTCAGAGAACTGAGGCTTGCCCATTTGGGAAGTGG	1260
Qy	1261	CCCACTGCCATTTGTGAGGGCCCAACCATCTTTGGGAGCTGTGGGAGCAAGGATCCC	1320
Db	1261	CCCACTGCCATTTGTGAGGGCCCAACCATCTTTGGGAGCTGTGGGAGCAAGGATCCC	1320
Qy	1321	CCAGTAACA 1329	
Db	1321	CCAGTAACA 1329	
RESULT 2			
US-10-114-104-108			
; Sequence 108, Application US/10114104			
; Publication No. US20030198647A1			
; GENERAL INFORMATION:			
APPLICANT: PERRON, HERVE			
BESEME, FREDERIC			
BEDIN, FREDERIC			
PARANHOS-BACCALA, GLAUCIA			
KOMURIAN-PRADEL, FLORENCE			
JOLIVET-REYNAUD, COLETTE			
MANDRAND, BERNARD			
GARSON, JEREMY			
TUKU, PHILIP			
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS			
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC			
THERAPEUTIC PURPOSES			
NUMBER OF SEQUENCES: 210			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: OLIFF & BERRIDGE, PLC			
STREET: P.O. BOX 1928			
CITY: ALEXANDRIA			
STATE: VA			
COUNTRY: USA			
ZIP: 22320			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/10/114.104			
FILING DATE: 03-Apr-2002			

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,847
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-10-114-104-108

Query Match 99.8%; Score 1326; DB 16; Length 1329;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAAATCGAAGAGCTTTAGACTGTCTAACCGCCAAAGAGGGGGAACCTGTTATTTT 60
DB 1 TCAAAATCGAAGAGCTTTAGACTGTCTAACCGCCAAAGAGGGGGAACCTGTTATTTT 60

QY 61 AGGGGAGAGATCGTGTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
DB 61 AGGGGAGAGATCGTGTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120

QY 121 AATTTGAGATCGAATATAATAGAGCAGAGAGACCTTCAAAACACTGCACTGGGGCT 180
DB 121 AATTTGAGATCGAATATAATAGAGCAGAGAGACCTTCAAAACACTGCACTGGGGCT 180

QY 181 CCTCAGCCAAATGAGATGCGCTGGAGCTCTCCCTTTAGGACCTCTAGGAGCTATAATTT 240
DB 181 CCTCAGCCAAATGAGATGCGCTGGAGCTCTCCCTTTAGGACCTCTAGGAGCTATAATTT 240

QY 241 TTTACTCCTTTTGGACCGCTGATCTTCACTTCTTGTAGTTGTTCTTCTTCCAGAAAT 300
DB 241 TTTACTCCTTTTGGACCGCTGATCTTCACTTCTTGTAGTTGTTCTTCTTCCAGAAAT 300

QY 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 360
DB 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 360

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTGATGTTAATGACATTTGA 420
DB 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTGATGTTAATGACATTTGA 420

QY 421 AGTCACCCCTCCGAGGAATCTCACTGACACACCCCTACTACTCCTCAATTCAGTAGG 480
DB 421 AGTCACCCCTCCGAGGAATCTCACTGACACACCCCTACTACTCCTCAATTCAGTAGG 480

QY 481 AAGCAGTTAGACAGTGTGACGCAACCTCCCAACAGTACTTGGGTTTTCTTGTGAGA 540
DB 481 AAGCAGTTAGACAGTGTGACGCAACCTCCCAACAGTACTTGGGTTTTCTTGTGAGA 540

QY 541 GGGTGGACTCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAGCCT 600
DB 541 GGGTGGACTCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAGCCT 600

QY 601 ANCTGGGAAGTGAACCCGATCCATCTTTAAACATGAGGCTTGAACCTAGCTACACCCG 660
DB 601 ANCTGGGAAGTGAACCCGATCCATCTTTAAACATGAGGCTTGAACCTAGCTACACCCG 660

QY 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
DB 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720

QY 721 AATCATCTATTGCTGTAGAGCAGACGCGGAAAGGAAGAGATTGGGATATAAATCAGGCA 780
DB 721 AATCATCTATTGCTGTAGAGCAGACGCGGAAAGGAAGAGATTGGGATATAAATCAGGCA 780

QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATTCATGTTATGGAGCTCTGTTTT 840
DB 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATTCATGTTATGGAGCTCTGTTTT 840

QY 841 CACTCTATTTCACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTGTTTTATGG 900
DB 841 CACTCTATTTCACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTGTTTTATGG 900

QY 901 CTCAGCTGAGCTTTTGTTCGCATCCACACTGCTGTTTGGCCACCGTCACAGACCCGCT 960
DB 901 CTCAGCTGAGCTTTTGTTCGCATCCACACTGCTGTTTGGCCACCGTCACAGACCCGCT 960

QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGTATCCAGCGAGGT 1020
DB 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGTATCCAGCGAGGT 1020

QY 1021 ACCCATTCGCCTCCCGATCAGGCTAAAGGCTTGGCAATGTTCTCTGTCATGGCTAAGTGCC 1080
DB 1021 ACCCATTCGCCTCCCGATCAGGCTAAAGGCTTGGCAATGTTCTCTGTCATGGCTAAGTGCC 1080

QY 1081 TGGGTTTTGCTTAATAGAACTGAACACTGGTCACTGGTTCATGGTTCTCTTCCATGAC 1140
DB 1081 TGGGTTTTGCTTAATAGAACTGAACACTGGTCACTGGTTCATGGTTCTCTTCCATGAC 1140

QY 1141 CCACGGCTTCTAATAGACTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA 1200
DB 1141 CCACGGCTTCTAATAGACTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA 1200

QY 1201 TCTGTGAGGCCAAGAACCCAGCTCAGAGAANGTGAAGGCTTGCCACCATTTGGGAAGTGG 1260
DB 1201 TCTGTGAGGCCAAGAACCCAGCTCAGAGAANGTGAAGGCTTGCCACCATTTGGGAAGTGG 1260

QY 1261 CCCACTGCCATTTGGTAGCGGCCCAACCACTTTGGAGCTGTGGGAGCAAGGATCCC 1320
DB 1261 CCCACTGCCATTTGGTAGCGGCCCAACCACTTTGGAGCTGTGGGAGCAAGGATCCC 1320

QY 1321 CCAGTAACA 1329
DB 1321 CCAGTAACA 1329

RESULT 3

US-09-873-367C-81
; Sequence 81, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2000-09-29
; PRIOR FILING DATE: 2000-09-29
; PRIOR FILING DATE: 2000-09-29
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093

; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-873-367C-81									
Query Match 77.9%; Score 1035.2; DB 10; Length 56093;									
Best Local Similarity 90.1%; Pred. No. 1.6e-310;									
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;									
Qy	1	TCAAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGAGGGGAACTGTTTATTTTT	60						
Db	37018	TCAAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTTT	37077						
Qy	61	AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGATCATTAAGTGAAGTTAAAGA	120						
Db	37078	AGGGGAAGATGCTGTTATATGTTAATCAATCGGAATCGTCACTGAGAAGTTAAAGA	37137						
Qy	121	AATTTGAGATCGAATATATATGAGCAGAGGACCTTCAAAACACCTGCACCCCTGGGGCT	180						
Db	37138	AATTCGAGATCGAATACACGCTAGAGCAGAGGAGCTTCGAACACCTGGACCTGGGGCT	37197						
Qy	181	CCTCAGCAATGGATGGCCCTGGACTCTCCCTTCTTAGGACTCTTAGCAGCTATAATAT	240						
Db	37198	CCTCAGCAATGGATGGCCCTGGATTTCTCCCTTCTTAGGACTCTTAGCAGCTATAATAT	37257						
Qy	241	TTTACTCCTCTTTGGACCCCTGATCTTCAACTCTTCTGTTAAGTTTGTCTCTCCAGAT	300						
Db	37258	GCTACTCCTCTTTGGACCCCTGATCTTTAACTCTCTTGTAACTTTGTCTCTCCAGAT	37317						
Qy	301	TGAAGCTGTAAGCTTACAAATAGTTTCTTCAATGGAAACCCAGATGCAGTCCATGACTAA	360						
Db	37318	CGNAGCTGTAAGCTA-----CAATGGAGCCCAAGATGCAGTCCAGACTAA	37365						
Qy	361	AATCTACCGTGAGACCCCTGGACCGGCTGTAGACTATGCTGTGATGTTAATGACATTGA	420						
Db	37366	GATCTACCGCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCAA	37425						
Qy	421	AGTACCCCTCCGAGGAATCTCAACTGACACACCCCTACTACACTCCAATTCAGTAGG	480						
Db	37426	AGGCACCCCTCCTGAGGAATCTCAGCTGCAACACCTCTACTAGCCCAATTCAGCAGG	37485						
Qy	481	AAGCAGTTAGAGCAGTTGTTCAGCCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540						
Db	37486	AAGCAGTTAGAGCGTCTGTCGCCCAACCTCCCAACAGCACTTAGTGTTCCTGTTGAGA	37545						
Qy	541	GGGTGGACTGAGACAGGACTAGCTGGAATTTCTTAGCTGACTAAGAAATCCCAAGCCT	600						
Db	37546	TGGGGGACTGAGACAGGACTAGCTGGAATTTCTTAGCTGACTAAGAAATCCCTAAGCCT	37605						
Qy	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	660						
Db	37606	AGCTGGGAAGGTGACCAATCCACTTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG	37665						
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCMAAAACAGGAGGTAAAGCAATAGCC	720						
Db	37666	ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCMAAGACAGGAGTAAAGAAATAGCC	37725						
Qy	721	AATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATGGATATAAATCAGGCA	780						
Db	37726	AATCATCTATTGCTGAGAGCAGCAGGAGGGAACAATGATCGGGATATAAACCAAGTC	37785						
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	840						
Db	37786	TTCAAGCCGGAACAGCAACCCCTTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTT	37845						
Qy	841	CACCTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTATGG	900						
Db	37846	CATGCTATTTCATCTATTAAATCTTGCACTGCACTCTTCTGGTCCATGTTTCTTACGG	37905						
Qy	901	CTAAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTTTCGACCGCTCAGACCCGCT	960						
Db	37906	CTTGAAGCTGAGCTTTGCTGCGCATCCACCACTGCTGTTTTCGCGCACCGCAGACCCG	37965						
Qy	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCTGCTGCTCTGATCCAGCGAGGT	1020						
RESULT 4									
US-10-087-192-910									
; Sequence 910, Application US/10087192									
; Publication No. US20020182586A1									
; GENERAL INFORMATION:									
; APPLICANT: Morris, David W.									
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR									
; FILE OF INVENTION: CANCER									
; FILE REFERENCE: 529452000122									
; CURRENT APPLICATION NUMBER: US/10/087,192									
; CURRENT FILING DATE: 2002-03-01									
; PRIOR APPLICATION NUMBER: US 09/747,377									
; PRIOR FILING DATE: 2000-12-22									
; PRIOR APPLICATION NUMBER: US 09/798,586									
; PRIOR FILING DATE: 2001-03-02									
; NUMBER OF SEQ ID NOS: 2059									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 910									
; LENGTH: 21646									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (1)...(21646)									
; OTHER INFORMATION: n = A,T,C or G									
US-10-087-192-910									
Query Match 60.1%; Score 798.2; DB 13; Length 21646;									
Best Local Similarity 87.1%; Pred. No. 7.6e-237;									
Matches 908; Conservative 0; Mismatches 106; Indels 28; Gaps 2;									
Qy	257	CCCTGTATCTTCAACTTCTTGTGTTAAGTTTGTCTTCTCCAGAAATGAAGCTGTTAAAGCTA	316						
Db	19571	CCCTGTATCTTAACTCTTGTGTTAAGTTTGTCTTCTCCAGAAATGAAGCTGTTAAACTA	19630						
Qy	317	CAATAGTTCTTCAATGGAACCCAGATCAGTCCATGCTAAATCTACCGTGGACCC	376						
Db	19631	CAATAGTTCTTCAATGGAACCCAGATCAGTCCATGCTAAATCTACCGTGGACCC	19690						
Qy	377	CTGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCTCCCGAG	436						
Db	19691	CTGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCTCCCGAG	19750						
Qy	437	GAATCTCAACTGCACAAACCCCTACTACATCCCAATTCAGTAGGAAGCAGTTAGAGCAGT	496						
Db	19751	GAATCTCAACTGTGCAACCCCTACTATGCCCAATTCAGCAGGAAGCAGTTACAGCGGT	19810						
Qy	497	TGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTTGGAGGGTGGACTGAGAGAC	556						
Db	19811	CATCAGCAACCTCCCAACAGCACTTTGGGTTTCTCTTGGAGGGGCGCACTGAGAGAC	19870						
Qy	557	AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCTTANCTGGGAAGGTGACC	616						

Db	19871	AAGACTAGCTGATTTCCCTTAGCGCAGTAAGAATCCCTAAGCGTAGCTGGGAAGGTGACT	19930
Qy	617	GCATCCATCTTTAAACATGGGGCTTGCACATTAGCTCACACCCGACCAATCAGAGAGCTC	676
Db	19931	GCATCCACCTTTAAACATGGGGCTTGCACATTAGCTCACACCCGACCAATCAGAGAGCTC	19990
Qy	677	ACTAAAATCCTAATCATAGGCCAAAAACAGGAGGTAAAGCAATAAGCCAATCATCTATTGGCTG	736
Db	19991	ACTAAAATCCTAATTTAGGCNAAAAACAGGAGGTAAAGCAATAAGCCAATCACCTTTGGCTG	20050
Qy	737	AGAGCACAGCGGGAAGGACAAGGATTTGGGATATAAACTCAGGCAATTCAAGCCGACAAAG	796
Db	20051	AGAGCACAGCGAGGAGCAAGGATTCGGGATATAAACCCGGGCAATTCAAGCCGCAATGG	20110
Qy	797	C AACCCCCCTTTGGGTCCCCTCCATTTGATATGGGAGCTCTGTTTTTCACTCTATTTTCACTCT	856
Db	20111	CAACCCCTTTGGGTCCCCTCCCTTTGATATGGGAGCTCTGT-----TTTCACTCT	20160
Qy	857	ATTAAATCATGCAACTGCACTCTTGTTGGTCCGCTGTTTTTATATGGCTCAAGCTGAGCTTTT	916
Db	20161	ATTAAATCTTGC AACTGCACTCTTCTGGTCCCAATGTTTACGGCTCGAGCTGAGCTTTC	20220
Qy	917	GTTTCGCATCCACCCTGCTGTTTGGCACCGTCACAGA-----CCCCG	958
Db	20221	GTCTCCCGTCCACCATGCTGTGTTGGCACTGTCGAGACCTGCCGCTGACTCCCATCTG	20280
Qy	959	CTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCA CTGTGCTCTGTATCCAGGAG	1018
Db	20281	CTGCTGACTCCCATCCCTCCAGATCCGGCAGGCTGTGCTGTGCTCTGACCCAGTGAG	20340
Qy	1019	GTACCAATGGCACTCCCGATCAGGCTAAAGGCTTGCCATTTGTTCTCGATGGCTAAGTG	1078
Db	20341	ACTCCCATTGCCACTCCCGATCGTGTCAAAGGCTTCCCATTTGTTCTGCAGGCTAAGTG	20400
Qy	1079	CTTGGGTTTGCTTAATAGNACTGACACTGGCTCACTGGGTTCCATGTTCTCTTCCATG	1138
Db	20401	CCTGGGTTGCTTAATCAGAGCTGAACACTAGTCACTGGGTTCCACAGTTCTCTTCCGCTG	20460
Qy	1139	ACCCAGGCTTCTAATAGAGCTATAACACTCAACCGCATGGGCCCAAGATTTCCATTCCTTGG	1198
Db	20461	ACCCAGGCTTCTAATAGAGCTATAACACTCACTCGCATGGGCCCAAGATTTCCATTCCTTGG	20520
Qy	1199	TATCTGTGAGGCCCAAGAACCCCAAGTTCAGAGAANGTGAGGCTTGCCACCATTTGGGAAGT	1258
Db	20521	AATCCGTGAGGCCCAAGAACCCCAAGTCAGAGAACA CAAGGCTTGCCACCATCTTGGGAAGC	20580
Qy	1259	GGGCCACTGCCATTTTGGGTAGC	1280
Db	20581	GGCTCACCACTCTTTGGGAGC	20602

[illegible]

Query Match	58.6%;	Score 778.8;	DB 17;	Length 1393;
Best Local Similarity	87.2%;	Pred. No. 1.9e-231;		
Matches 904;	Conservative	0;	Mismatches 112;	Indels 21; Gaps 4
Qy	302	GAAGCTGTAAAGCTACAAAATAGTTCTTCAAATGAAACCCACAGATGCAGTCCATGACTAAA	361	
Db	1	GAAGCTGTAAAACTACTAAATGTGTTCTTCAAAGGAGCCCGAGATGCAGTCCATGACTAAG	60	
Qy	362	ATCTACCGTGGACCCCTGGACCGCGCTGTAGACTATGCTCTGTATGTTAAATGACATTGAA	421	
Db	61	ATCTACACACAGACCCCTGGACCGCGCTGTAGCCCATGCAACCGATGTTAATGATATCGAN	120	
Qy	422	GTCAACCCCTCCCGAGGAAATCTCAACTGCGACAAACCCCTACTACACTCCAAATTCAGTAGGA	481	
Db	121	AGCACTCC-CCCAAGGAAATTTCACTGCGACAAACCCCTACTACACCCCAATTCAGCAGGA	179	
Qy	482	AGCAGTTAGAGCAGTTGTGACGCCAACTCCCCAAACAGTACTTGGGTTTTCTGTTGAGAG	541	
Db	180	AGCAGTTAGAGCGGTGCTGACGCCAACCTCCCGAGCAGCACTTGGGTTTTCTGTTGAGAG	239	
Qy	542	GGTGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCCTA	601	

```

Db 240 CCGGACCTGAGACAGACAGACTAGCTGGATTTCTTAGCTGACTAAGAATCCCTAAGCCTA 299
Qy 602 NCTGGGAAGTGACCGCATCCATCTTTAAACATGCGGCTTGCAACTTAGCTCACAACCGGA 661
Db 300 GCTGGGAAGTGACCGCTCTACCTTTAAACACAGCGGCTTGCAACTTAGCTCACAACCGCA 359
Qy 662 CCAATC-----AGAGAGCTCACTAAATGCTTAATCAGGCACAAAACAGAGGTAAG 712
Db 360 CCAATCAGGTAGTAAAGAGAGCTCACTAAATGCTTAATTAGGCACAAAACAGAGGTAAG 419
Qy 713 CAATAGCCAATCATCTATTGCTGCTGAGACACAGCGGGAAGGACAGAGGATGGGATATAA 772
Db 420 AAATAGCCATCATCTATCGCTGAGACACAGAGGAGGACATGATCGGATATAA 479
Qy 773 CTCAGGCATTAAGCAGCAACAGACCCCTTTGGGTCCCTCCCAATGTATGGGAGC 832
Db 480 CCCAAGCATTCGAGCAGCAACGGCTGCCCTTTGTGCTCCCTTTGTATGGGAGC 539
Qy 833 TCTGTTTTCACCTCACTATTTCACCTATTAAATCATGCAACTGCACTCTTCTGGTCGGTGT 892
Db 540 TCTGT-----TTTCACTCTATTAAATCTTGCMACTGCACTCTTCTGGTCGTGT 589
Qy 893 TTTTATGGCTCAAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTTGCACCGTCA 952
Db 590 TGTACGGTTTGAGCTGAGCTTTCGCTCGCGTCCACCACTGCTGTTTGC CGCATCGGA 649
Qy 953 GACCGCTGCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCACACTGTCTCTGATCC 1012
Db 650 GACCTGCGCTGACTTCCATCCCTCCGATCTGGCAGGGTGTTCA-TGTGCTCCTGATCC 708
Qy 1013 AGCGAGTACCATTGCCACTCCCGATCAGCTTAAGGCTTGCCATTTCTCTGCTGATGC 1072
Db 709 AGAGAGGACCCATTGCCATTCCTGATTGGGCTTAAGGCTTGCCATTTCTCTGCGANG 768
Qy 1073 TAAAGTCCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCTCACTGGGTTCCATGGTTCTCT 1132
Db 769 TAAGTCCCGGGTTTCATCTTAATCGAGCTGAACACTAGTCTGCTGGTTCCAGATTCTCT 828
Qy 1133 TCCATGACCCAGGGTTCTTAATAGAGCTTAACACTCAACCGATGCGCCAGATTCATTT 1192
Db 829 TCCGTGACCCAGCACTTCTTAATAGAGCTTAACACTCAACCGACCGGCCAACAGATTCCATT 888
Qy 1193 CCTTGGTATCTGTAGGCAAGAACCCAGAGTCAAGAGAACTGAGGCTTGCCACCATTTG 1252
Db 889 CTTTGAATTCGTGAGGCAAGAACCCAGGTCAGAGAACTAGAGGCTTGCCACCATTTT 948
Qy 1253 GGAAGTGCCCACTGCCATTTTGGTGTAGCGGCCCAACCACTCTTGGGAGCTGTGGGAGCA 1312
Db 949 GGAAGTGCCCTGCGCCATTTTGGAAAGTGCGCTGCCACCATCTTGGGAGCTCTGGGAGCA 1008
Qy 1313 AGGATCCCCAGTAACA 1329
Db 1009 AGGACCCCTCTGTAACA 1025

RESULT 6
US-10-637-565-18
; Sequence 18, Application US/10637565
; Publication No. US2004004381A1
; GENERAL INFORMATION:
; APPLICANT: PERRON Herve
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: KOMURIAN-PRADEL, Florence
; TITLE OF INVENTION: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES
; TITLE OF INVENTION: METHODS FOR DETECTING MSRV-1 RETROVIRUS
; FILE REFERENCE: 110257
; CURRENT APPLICATION NUMBER: US/10/637,565
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US/09/890,340
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/IB00/00159
; PRIOR FILING DATE: 2000-02-15

```

```

; PRIOR APPLICATION NUMBER: EP 99420041.8
; PRIOR FILING DATE: 1999-02-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 2030
; TYPE: DNA
; ORGANISM: MSRV-1 retrovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1626)
US-10-637-565-18

Query Match 58.3%; Score 775.4; DB 17; Length 2030;
Best Local Similarity 92.8%; Pred. No. 2.7e-230;
Matches 812; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 1 TCATAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCAAAAGAGGGGGAACCTGTTTATTTT 60
Db 1140 TCATAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCAAAAGAGGGGGAACCTGTTTATTTT 1199
Qy 61 AGGGAAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTAATCTGAGAAAGTTAAAGA 120
Db 1200 AGGAGAAAGACGCTGTTATTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGA 1259
Qy 121 AATTTGAGATCGAATATAATGTAGACAGAGGACCTTCMAAACACTGCACCTCGGGCCT 180
Db 1260 AATTCGAGATCGAATAAATGTAGACAGAGGACCTTCMAAACACTGCACCTCGGGCCT 1319
Qy 181 CCTCAGCAATGGATGCCCTGGGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 1320 CCTCAGCAATGGATGCCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 1379
Qy 241 TTTACTCTCTTTGGACCCCTGTAATCTTCAACTCTCTGTTAAAGTTGTTCTTCCAGAAAT 300
Db 1380 GTTACTCTCTCTTTGGACCCCTGTAATCTTAACTCTCTGTTAAAGTTGTTCTTCCAGAAAT 1439
Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCACTGCATGCATGACTAA 360
Db 1440 TGAAGCTGTAAGCTACAGATGTTCTTCAATGGAACCCAGATGCACTGCATGCATGACTAA 1499
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTTGA 420
Db 1500 GATCCACCGTGGACCCCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACTTGA 1559
Qy 421 AGTCACCCCTCCGAGAAATCTCAACTGCACACCCCTACTACACTTCCAATTCAGTAGG 480
Db 1560 AGGCACCCCTCCGAGAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCAGCGGG 1619
Qy 481 AAGCAGTTAGAGCAGTTGTCAGCCCACTCTCCCAACAGTACTTGGGTTTTCCTGTTGAGA 540
Db 1620 AAGCAGTTAGAGCGGTGTCAGCCCACTCTCCCAACAGCAGCTTGGGTTTTCCTGTTGAGA 1679
Qy 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTTCTTAGGCTGACTAAGAATCCCAAGCCT 600
Db 1680 GGGGGGACTGAGAGACAGGACTAGCTGGATTTTCTTAGGCTGACTAAGAATCCCAAGCCT 1739
Qy 601 ANCTGGGAAGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACAACCCG 660
Db 1740 AGCTGGGAAGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACAACCCG 1799
Qy 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCACAAAACAGGAGGTAAAGCAATAGCC 720
Db 1800 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCACAAAACAGGAGGTAAAGCAATAGCC 1859
Qy 721 AATCATCTATTGCTGAGACACAGCGGGAAGCAAGGATTTGGGATATAAATCAGGCA 780
Db 1860 AATCATCTATTGCTGAGACACAGCGGGAAGCAAGGATTCGGGATATAAATCAGGCA 1919
Qy 781 TTCNAGCCAGCAACAGCAACCCCTTTGGTCCCTCTCCCATTTGATGGAGCTCTGTTTT 840
Db 1920 TTCGAGCCGCAACGGCAACCCCTTTGGTCCCTCTCCCTTTGATGGGCTCTGTTTT 1979

```

Qy	841	CAC	TCT	ATT	CAC	TCT	ATT	TAA	AAT	CAT	GCA	CTG	CA	875
D _b	1980	CAC <th>TCT</th> <th>ATT</th> <th>CAC</th> <th>TCT</th> <th>ATT</th> <th>TAA</th> <th>AAT</th> <th>CAT</th> <th>GCA</th> <th>CTG</th> <th>CA</th> <th>2014</th>	TCT	ATT	CAC	TCT	ATT	TAA	AAT	CAT	GCA	CTG	CA	2014

RESULT 7

```

RESULT 7
US-10-198-846-9936
; Sequence 9936, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9936
; LENGTH: 7974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-9936

```

	Query Match	Best Local Similarity	Mismatches	Conservative	0;	Mismatches	Indels	29;	Gaps	6;
Qy	257	CCCTGTATCTTCAACTTCCTTGGTTAAAGTTGTCTCTTCCAGAATGTAAGCTGTAAAGCTA	316							
nb	6409	CCCTGTATCTTCAACTTCCTTGGTTAAAGTTGTCTCTTCCAGAATGTAAGCTGTAAAGCTA	6468							

QY	257	CCCTGTATCTTCAACTTCCCTTGTGTTAAAGTCTCTCTTCCAGAAATGGAAGCTGTAAAGCTA	316
Db	6409	CCCTGTATCTTAAACCTCCCTTGTGTAAAGTCTCTCTTCCAGAAATCAAAGCTGTAAAACTA	6468
QY	317	CAAAATAGTCTTCAAATGGAAACCCAGATGCAGTCCATGAC - TAAAAATCTACCGTGGACC	375
Db	6469	CAAATCGTTCTTCAAATGGAGCCCAAGATGCAGTCCATGACTTAAAGATCTTACCATGGACC	6528
QY	376	CCTGGACGGCCTGCTAGACTATGCTCTGATGTTTAATGA CATTGAAGTCAACCCCTCCCGA	435
Db	6529	CCTGGACGAGCTGCTAGCCCATGCTCTGATGTTTAATGA CATCGAAGACACCCCTCCAGA	6588
QY	436	GGAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGGAGCAGTTAGACGAG	495
Db	6589	GGAATCTCAACTGCAACACCCCTACTATGCCCTGATTTCAGCAGGAAGTAGTTAGAGTGG	6648
QY	496	TTGTGACGCAACCTCCCCAAACAGTACTTTGGGTGTTTTCTGTTTGACAGAGGTGGACTGAGAGA	555
Db	6649	TTGTGCGGCAACCTCCCCAAACAGCATTGAGTTTTCTGTTTGGAGGGGGAGCTGACAGA	6708
QY	556	CAGGACTAGCTGGATTTTCCCTA - -----GGCTGACTAAGAAATCCCNAAAGCCTTANCTGGGA	608
Db	6709	CAGGACTAGCTGGGATTTCCCTAGGGCCGACTTAAAGAATCACCACCTAAGCCTAGCTGGGA	6768
QY	609	AGGTGACCGCAT - CCAATCTTTAAACATGGGGCTTGCAACTTTAGTCTCACACCCGACCAATC	667
Db	6769	AGGTGACCGCGTCCACCTTTAAACACGGGGCTTGCAACTTTAGTCTCACACCCGACCAATC	6828
QY	668	-----AGAGTACTCACTTAAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAG	718
Db	6829	AGGTAGCAAGAGAGCTCGTTTAAATGATTAATTAGGCAAAAACAGGAGGTAAAGAAATAG	6888
QY	719	CCAATCATC - TATTGCTCTGAGACGACAGCGGGGAAGGACAAGGATTTGGGATATAAACTCAG	777
Db	6889	CCAATCATCTTATCACCTGAGACACAATGGGAAGGACAATGATCGGGATATAAAACCCAG	6948
QY	778	GCATTTCAAGCCAGCAACAGCAACCCCTTTTGGTCCCTCCCATTTGATTTGGGAGCTCTGT	837
Db	6949	GCATTTCCAGTTTGGCAACTGCTACCTCTTTGGGTCCCTCCCTTTGATATGGAAGCTCTGT	7008

Qy	838	TTTCACTCTATTTCACCTCTCTATTAAATCATGCAACTGCACCTCTCTCTGGTCCGTGTTTTTA	897
Db	7009	-----TTTTCACCTCTATTAAATCTTGCTCTTCGACGCTCTCTGGTCCATATTGTTA	7058
Qy	898	TGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTTGCCACCGTCACAGACC	957
Db	7059	CGGCTCCAGCTGAGCTTTTGTCTCGCGCTCCACCCCTGTGTTTGGCCGCTGTCGACAGACC	7118
Qy	958	GCTGCTGACTTCACATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCACGGA	1017
Db	7119	GCGCGTAGCTTCATCCCTCCGGATCCAGCAGGCTGTCTGCTGTCTCTGATGTCAGGA	7178
Qy	1018	GGTACCCATTGCCACTCCGATCAGGCTTAAGGCTTGGCCATGTTGCTGCATGGCTAAAGT	1077
Db	7179	GGCGCCCATTTCCACTCCTGATCGGACTGAAGGCTTGCCATTGTTCTGTCACGGCTAAGA	7238
Qy	1078	GCCTGGGTTTGTCTTAATAGAACTGAACACTGCTGCTCACTGGGTTCCATGGTTCTCTCCAT	1137
Db	7239	GCCTGGGTTGCTCTTAATCAGCTGAACACTTAGTCACTGGGTTCCACGGTTCTCTCCGT	7298
Qy	1138	GACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATCCATTCTCTTG	1197
Db	7299	GACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATCCATTCTCTTG	7358
Qy	1198	GTATCTGTGAGGCCAAGAACCCGAGTCAAGAAAGTGTAGGCTTGCCACCATTTGGGAAG	1257
Db	7359	GAATCCGTGAGGCCAAGAACCCGAGTCAAGAAACACAGAGGCTTGCCACCATTTGGGAAG	7418
Qy	1258	TGGGCCACTGCCATTTTGGTAGCGGCCACACCATCTTGGGAGCTGTGGAGCAAGGAT	1317
Db	7419	TGGGCTGCCACCATCTTTGGAAGCGGCTTGCCACCATCTTGGGAGCTGTGGAGCAAGGAC	7478
Qy	1318	CCCCAGTAACA	1329
Db	7479	CCCCGGTAACA	7490

RESULT 8

```

US-10-416-642-4
; Sequence 4, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: ARVIZU, Chandra
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0842 PCT
; CURRENT APPLICATION NUMBER: US/10/416,642
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/249,407
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7477736CB1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1995
; OTHER INFORMATION: a, t, c, g, or other
; US-10-416-642-4

```

Query Match 57.7%; Score 766.8; DB 17; Length 2074;
Best Local Similarity 93.0%; Pred. No. 1.3e-227;
Matches 812; Conservative 0; Mismatches 60; Indels 1; Gaps 1
Qy 1 TCRAAATCGAAGACCTTTAGACTTGTAAACGCCAAGAGGGGGAACTGTTTATTTT 60

QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCATTTGATGAGAGCTCTGTGTTT 840
DB 2835 TTCAGCCGCGCAACGCAACCCCTTTGGTCCCTCCCTTTGATGAGAGCTCTGTGTTT 2894
QY 841 CACTCTATTTTCACTCTATTAAATCATGCAACTGCA 875
DB 2895 CATGCTATTTTCACTCTATTAAATCTTGCAACTGCA 2929

RESULT 10
US-10-016-249-3
; Sequence 3, Application US/10016249
; Publication No. US20030100053A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 606B.A1172A
; CURRENT APPLICATION NUMBER: US/10/016,249
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US/09/175,928
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-249-3

Query Match 55.1%; Score 731.8; DB 15; Length 2946;
Best Local Similarity 90.7%; Pred. No. 1.3e-216;
Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTGTTATTTT 60
DB 2067 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGAACTGTTATTTT 2126
QY 61 AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
DB 2127 AGGGGAAGAAATGCTGTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 2186
QY 121 AATTTGAGATCGAATATATAGACAGAGGACCTTCAAAACACTGCACCTGGGGCCT 180
DB 2187 AATTCGAGATCGAATACACGATAGACAGAGGAGCTTCAAAACACTGGAGCCCTGGGGCCT 2246
QY 181 CCTCAGCCAAATGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 240
DB 2247 CCTCAGCCAAATGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 2306
QY 241 TTTACTCTCTTTGGACCCCTGATATCTTCAACTCTCTTTGATTTGTTGTTCTCTTCCAGAT 300
DB 2307 GCTACTCTCTTTGGACCCCTGATATCTTCAACTCTCTTTGATTTGTTGTTCTCTTCCAGAT 2366
QY 301 TGAAGCTGTAAGACTCAAAATAGTTCTTCAATGGAACCCCAAGATCGAGTCCATGACTAA 360
DB 2367 CGAAGCTGTAAGACTCAAAATAGTTCTTCAATGGAACCCCAAGATCGAGTCCATGACTAA 2414
QY 361 AATCTACCTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGTA 420
DB 2415 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 2474
QY 421 AGTCACCCCTCCCGAGAAATCTCAAATGACAAACCCCTACTACATCTCCAAATTCAGTAG 480
DB 2475 AGGCACCCCTCCCGAGAAATCTCAGCTGCACAAACCTCTACTACGCCCAATTCAGCAGG 2534

QY 481 AAGCAGTTAGAGCAGTGTGTAGCCAAACCTCCCAACAGTACTTGGGTTTTCTGTGTGAGA 540
DB 2535 AAGCAGTTAGAGCAGTGTGTAGCCAAACCTCCCAACAGTACTTGGGTTTTCTGTGTGAGA 2594
QY 541 GGGTGGACTGAGAGCAGGAGCTAGCTGATTTCTTAGGCTGACTAAGATCCCAAGCCT 600
DB 2595 TGGGGGACTGAGAGCAGGAGCTAGCTGATTTCTTAGGCTGACTAAGATCCCAAGCCT 2654
QY 601 ANCTGGGAAGGTGACCGCATCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCG 660
DB 2655 AGCTGGGAAGGTGACCATCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCTG 2714
QY 661 ACCAATCAGAGAGCTCACTAAATCTAATCAGGCAAAAAAGAGGTTAAAGCAATAGCC 720
DB 2715 ACCAATCAGAGAGCTCACTAAATCTAATCAGGCAAAAAAGAGGTTAAAGCAATAGCC 2774
QY 721 AATCATCTATTGCTGAGAGCAGGAGGAGGAGGATTTGGGATATAAACTCAGGCA 780
DB 2775 AATCATCTATTGCTGAGAGCAGGAGGAGGAGGATTTGGGATATAAACTCAGGCA 2834
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGCTCCCTCCCATTTGTATGGAGCTCTGTTTT 840
DB 2835 TTCAAGCCAGCAACAGCAACCCCTTTGGGCTCCCTCCCATTTGTATGGAGCTCTGTTTT 2894
QY 841 CACTCTATTTTCACTCTATTAAATCATGCAACTGCA 875
DB 2895 CATGCTATTTTCACTCTATTAAATCTTGCAACTGCA 2929

RESULT 11
US-09-902-535-1
; Sequence 1, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: Mi, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preeclampsia and gestational trophoblast
; TITLE OF INVENTION: disorders
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930) ... (2546)
US-09-902-535-1

Query Match 55.0%; Score 730.8; DB 9; Length 2930;
Best Local Similarity 90.7%; Pred. No. 2.7e-216;
Matches 793; Conservative 0; Mismatches 69; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTGTTATTTT 60
DB 2069 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGAACTGTTATTTT 2128
QY 61 AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
DB 2129 AGGGGAAGAAATGCTGTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 2188
QY 121 AATTTGAGATCGAATATAATGAGCAGAGGAGCTTCAAAACACTGCACCTGGGGCCT 180
DB 2189 AATTCGAGATCGAATACACGATAGACAGAGGAGCTTCAAAACACTGGAGCCCTGGGGCCT 2248
QY 181 CCTCAGCCAAATGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 240

Db 2249 CCTAGCCAAATGATCCCTGGATTCCTCCCTCTTAGACCTCTAGCAGCTATAATTT 2308
Qy 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCCTTTAAGTTTGTCTTTCAGAAAT 300
Db 2309 GCTACTCTCTTTGGACCTGATCTTAAACCTCTTAACTTTGTCTCTTCCAGAA 2368
Qy 301 TGAAGCTTGAAGCTTACAATAGTTCTTCAATGGAACCCAGATGCAATCCATGACTAA 360
Db 2369 CGAAGCTGTAATACTA-----CAATGGAGCCCAAGATGCAAGTCCAAGACTAA 2416
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db 2417 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCA 2476
Qy 421 AGTACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTACACTCAATTCAGTAGG 480
Db 2477 AGGCACCCCTCTGAGGAAATCTCAGCTGCACAACTCTACTAGCCCAATTCAGCAGG 2536
Qy 481 AAGCAGTTAGACAGTTCTCAGCCCAACCTCCCAACAGCAGCTTGGGTTTCTCTGTGAGA 540
Db 2537 AAGCAGTTAGACGCGTCTGCGCCCAACCTCCCAACAGCAGCTTAGGTTTCTCTGTGAGA 2596
Qy 541 GGGTGAAGCTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGCAATCCNAAAGCCT 600
Db 2597 TGGGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGCAATCCCTAAGCCT 2656
Qy 601 ANCTGGGAAGTGACCGCATCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 2657 AGCTGGGAAGTGACCATCATCCACTTTAAACACCGGGCTTGCAACTTAGCTCACACCTG 2716
Qy 661 ACCAATCAGAGCTCACTAAATGCTAATCAGCAAAACAGGAGTAAAGCAATAGCC 720
Db 2717 ACCAATCAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGTAAAGCAATAGCC 2776
Qy 721 AATCATCTATTGCTCTGAGACACAGCGGAAGGACAAGGATTTGGGATATAAATCAGGCA 780
Db 2777 AATCATCTATTGCTCTGAGACACAGCAGGAGGACAATGATCGGATATAAATCAGGCA 2836
Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATGTATGGAGGCTCTGTTTT 840
Db 2837 TTCGAGCGGCAACGCAACCCCTTTGGGTCCCTCTTTGTATGGAGCTCTGTTTT 2896
Qy 841 CACTCTATTTCACCTATTAAATCATGCAACTGC 874
Db 2897 CATGCTATTTCACCTATTAAATCTTGCAACTGC 2930

RESULT 12

US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauclia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; FILE OF INVENTION: 110048
; FILE OF INVENTION: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: misc feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-30

Query Match 54.4%; Score 722.6; DB 17; Length 7582;
Best Local Similarity 88.5%; Pred. No. 1.6e-213;
Matches 774; Conservative 19; Mismatches 70; Indels 12; Gaps 1;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTCTTAAACGCAAGAGGGGGAACCTGTTATTTT 60
Db 6720 TCRAAATCGAAGAGCTTTAGACTTGTCTTAAACGCTGAGAGAGGGGGAACCTGTTATTTT 6779
Qy 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCAATCTGGAAGTTAAAGA 120
Db 6780 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCTGCTACTGAGAAGTTAAGA 6839
Qy 121 AATTGAGATCGAATAATAATGTAGACGAGAGGACCTTCAAAACACTGCACCTCGGGCCT 180
Db 6840 AATTCGAGATCGAATAACAACGTAAGAGAGAGCTTGGAAACACTGCACCTCGGGCCT 6899
Qy 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTTTAGGACCTCTAGCAGTATAATTT 240
Db 6900 CCTCAGCAATGGATGCCCTGGACTCTCCCTTTTAGGACCTCTAGCAGTATAATTT 6959
Qy 241 TTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTTGAAGTTTGTCTCTCCAGAA 300
Db 6960 GCTACTCTCTTTGGACCCCTGATCTTTRACCTCTTGTAACTTTGTCTCTCCAGAA 7019
Qy 301 TGAAGCTGTAAGAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAAGTCAATGACTAA 360
Db 7020 CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATGCAAGTCAAGACTAA 7067
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db 7068 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCA 7127

QY 421 AGTACCCCTCCCGAGGAAATCTCAATGTCACAAACCCCTACTACACTCAATTCAGTAGG 480
DB 7128 AGGCACCCCTCTCTAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 7187
QY 481 AAGCAGTTAGACAGCTTCTCAGCCAACTCCCAACAGTACTTGGTTTCTCTGTTGAGA 540
DB 7188 AAGCAGTTAGACAGCTTCTCAGCCAACTCCCAACAGCAGTACTTGGTTTCTCTGTTGAGA 7247
QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGCTAAGAAATCCCAAGCCT 600
DB 7248 TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGCTAAGAAATCCCAAGCCT 7307
QY 601 ANCTGGGAAGTGACCGCATCATCTTTAAACATGGGGCTTGCACACTAGCTCACAACCG 660
DB 7308 AGSTGGGAAGTGACCAATCCACTTTAAACACCGGGCTTGCACACTAGCTCACAACCTG 7367
QY 661 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
DB 7368 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 7427
QY 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
DB 7428 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 7487
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 840
DB 7488 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 7547
QY 841 CACTCTATTTCATCTATTAAATCATGCAACTGCA 875
DB 7548 CATGCTATTTCATCTATTAAATCTTGCARCTGCR 7582

RESULT 13
US-10-632-793-25
; Sequence 25, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauca
; APPLICANT: MALLET, Cecile
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; FILE REFERENCE: 110048
; CURRENT FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2001-10-22
; PRIOR FILING DATE: 2000-01-21
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 25
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-25
Query Match 54.3%; Score 721.2; DB 17; Length 1136;
Best Local Similarity 90.0%; Pred. No. 1.6e-213;
Matches 787; Conservative 0; Mismatches 75; Indels 12; Gaps 1;
QY 1 TCAAAATCGAGAGCTTACACTTGTAAACCGCAAGAGGGGACCTGTTTATTTT.60
DB 254 TCAAAATCGAGAGCTTACACTTGTAAACCGCTGAAGAGGGGACCTGTTTATTTT 313
QY 61 AGGGGAAGATGCTGTTAGTATGTAATCAATCTGGAATCATTACTAGAGAAAGTAAAGA 120
DB 314 AGGGGAAGATGCTGTTATGTTAATCAATCCGGNATCGTCACTAGAGAAAGTAAAGA 373

QY 121 AATTGAGATCGAATATATGTAGAGCAGAGGACCTTCAAAACACTGCACCTCGGGCCT 180
DB 374 AATTGAGATCGAATATATGTAGAGCAGAGGAGCTTCAAAACACTGCACCTCGGGCCT 433
QY 181 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTTAGGACCTCTAGCAGCTATATATTT 240
DB 434 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTTAGGACCTCTAGCAGCTATATATTT 493
QY 241 TTTACTCTCTTTGAGACCTGATCTTCAACTCTCTTGTAAAGTTTGTCTTCCAGAAT 300
DB 494 GCTACTCTCTTTGAGACCTGATCTTAAACCTCTCTTGTAAAGTTTGTCTTCCAGAAT 553
QY 301 TGAAGCTGTAAGACTTCAAAATAGTTTCTTCAAAATGGAACCCAGATGCACTGCACTAA 360
DB 554 CGAAGCTGTAAGACTT-----CAAAATGGAACCCAGATGCACTGCACTAA 601
QY 361 AATCTACCGTGGACCCCTGGAGCCGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
DB 602 GATCTACCGCAGACCCCTGGAGCCGCTGCTAGCCACGATCTGATGTTAATGACATCAA 661
QY 421 AGTACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAATTCAGTAGG 480
DB 662 AGGCAACCCCTCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGCAGG 721
QY 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTTCCCAACAGTACTTGGGTTTTCTCTGTTGAGA 540
DB 722 AAGCAGTTAGAGCAGTTGTGAGCAACCTTCCCAACAGTACTTGGGTTTTCTCTGTTGAGA 781
QY 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
DB 782 TGGGGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 841
QY 601 ANCTGGGAAGTGACCGCATCATCTTTAAACATGGGGCTTGCACCTTACTGCACACCCG 660
DB 842 AGTGGGAAGTGACCAATCCACTTTAAACAGGGGCTTGCACCTTACTGCACACCTG 901
QY 661 ACCAATCAGAGACTCACTATAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
DB 902 ACCAATCAGAGACTCACTATAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 961
QY 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
DB 962 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATATAAACTCAGGCA 1021
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 840
DB 1022 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 1081
QY 841 CACTCTATTTCATCTATTAAATCATGCAACTGC 874
DB 1082 CATGCTATTTCATCTATTAAATCTTGCAGCTGC 1115

RESULT 14
US-10-632-793-26
; Sequence 26, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauca
; APPLICANT: MALLET, Cecile
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; FILE REFERENCE: 110048
; CURRENT FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2001-10-22
; PRIOR FILING DATE: 2000-01-21
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-26

Query Match      53.6%; Score 712.6; DB 17; Length 2782;
Best Local Similarity 89.4%; Pred. No. 1.2e-210;
Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;

QY 1 TC AAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTCTTTATTTTT 60
DB 1902 TCGAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAGAGGGGAACTCTTTATTTTT 1961

QY 61 AGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCAATTAAGTGAAGTTAAAGA 120
DB 1962 AGGGGAAGAATGCTGTATATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGA 2021

QY 121 AATTGGAGATCGAATATATGATAGAGCAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
DB 2022 AATTCCAGATCGAATATACACGTATAGCAGAGGAGCTTCGAAACACTGCAGCCTGGGGCT 2081

QY 181 CCTCAGCAATGGATGCCCTGACTCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 240
DB 2082 CCTCAGCGATGGATGCCCTGATCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 2141

QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTGTCTTCCAGAAAT 300
DB 2142 GCTACTCTCTTTGGACCTGTATCTTGTGACTCTCTTGTAACTTGTCTCTTCCAGAAAT 2201

QY 301 TGAAGCTGTAAGCTACAATAGTCTTCAATGGAAACCCAGATGCAATCCATGACTAA 360
DB 2202 CGAAGCTGTGAACCTA-----CNAATGGAGCCCAAGATGCAAGTCCAAGACTAA 2249

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 2250 GATCTACCGGATGAGACAGGACTAGCTGGATTTCTCTAGCCAGACTGATGTTAATGACATCA 2309

QY 421 AGTCACCCCTCCGAGGAAATCTCAATGTGCAACACCCCTACTACACTCCAATTCAGTAGG 480
DB 2310 AGGCACCCCTCTGAGGAAATCTCAGCTGCAACACTCTACTACGCCCAATTCAGCAGG 2369

QY 481 AAGCAGTTAGACAGTGTGACGCAACCTCCCAACAGTACTTGGGTTTTCTCTTTGAGA 540
DB 2370 AAGCAGTTAGACGCGTGTGCGCCAACTCCCAACAGCAGTCTAGGTTTTCTCTTTGAGA 2429

QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 600
DB 2430 TGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 2489

QY 601 ANCTGGGAAGGTGACCGCATCGATCTTTAAACATGCGGCTTGCAACTTAGCTCACACCG 660
DB 2490 AGGTGGGAAGGTGACCATCACCTTTAAACACGCGGCTTGCAACTTAGCTCACACCTG 2549

QY 661 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
DB 2550 ACCAATCAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAGCAATAGCC 2609

QY 721 AATCATCTATTGCTCTGAGAGCACAGCGGAAGGACAAGGATTTGGGATATAAATCAGGCA 780
DB 2610 AATCATTTATTGCTCTGAGAGCACAGCAGGAGGACAATGATCGGGATATAAATCAGGCA 2669

QY 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATCTGATGGAGCTCTGTTTT 840
DB 2670 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATCTGATGGAGCTCTGTTTT 2729

QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGCA 875
DB 2730 CATGCTATTTCACCTCTATTAAATCATGCAACTGCA 2764
```

RESULT 15

```
US-10-133-036-1
; Sequence 1, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Human endogenous retrovirus
US-10-133-036-1

Query Match      53.5%; Score 711; DB 17; Length 2782;
Best Local Similarity 89.3%; Pred. No. 3.9e-210;
Matches 781; Conservative 0; Mismatches 82; Indels 12; Gaps 1;

QY 1 TC AAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTCTTTATTTTT 60
DB 1902 TCGAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAGAGGGGAACTCTTTATTTTT 1961

QY 61 AGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCAATTAAGTGAAGTTAAAGA 120
DB 1962 AGGGGAAGAATGCTGTATATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTGAAGA 2021

QY 121 AATTGGAGATCGAATATATGATAGAGCAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
DB 2022 AATTCCAGATCGAATATACACGTATAGCAGAGGAGCTTCGAAACACTGCAGCCTGGGGCT 2081

QY 181 CCTCAGCAATGGATGCCCTGACTCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 240
DB 2082 CCTCAGCGATGGATGCCCTGATTTCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 2141

QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTGTCTTCCAGAAAT 300
DB 2142 GCTACTCTCTTTGGACCTGTATCTTGTGACTCTCTTGTAACTTGTCTCTTCCAGAAAT 2201

QY 301 TGAAGCTGTAAGCTACAATAGTCTTCAATGGAAACCCAGATGCAATCCATGACTAA 360
DB 2202 CGAAGCTGTGAACCTA-----CNAATGGAGCCCAAGATGCAAGTCCAAGACTAA 2249

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 2250 GATCTACCGGAAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCAA 2309

QY 421 AGTCACCCCTCCGAGGAAATCTCAATGTGCAACACCCCTACTACACTCCAATTCAGTAGG 480
DB 2310 AGGCACCCCTCTGAGGAAATCTCAGCTGCAACACTCTACTACGCCCAATTCAGCAGG 2369

QY 481 AAGCAGTTAGACAGTGTGACGCAACCTCCCAACAGTACTTGGGTTTTCTCTTTGAGA 540
DB 2370 AAGCAGTTAGACGCGTGTGCGCCAACTCCCAACAGCAGTCTAGGTTTTCTCTTTGAGA 2429

QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 600
DB 2430 TGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 2489

QY 601 ANCTGGGAAGGTGACCGCATCGATCTTTAAACATGCGGCTTGCAACTTAGCTCACACCG 660
DB 2490 AGGTGGGAAGGTGACCATCACCTTTAAACACGCGGCTTGCAACTTAGCTCACACCTG 2549

QY 661 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
DB 2550 ACCAATCAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAGCAATAGCC 2609

QY 721 AATCATCTATTGCTCTGAGAGCACAGCGGAAGGACAAGGATTTGGGATATAAATCAGGCA 780
DB 2610 AATCATTTATTGCTCTGAGAGCACAGCAGGAGGACAATGATCGGGATATAAATCAGGCA 2669

QY 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATCTGATGGAGCTCTGTTTT 840
DB 2670 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATCTGATGGAGCTCTGTTTT 2729

QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGCA 875
DB 2730 CATGCTATTTCACCTCTATTAAATCATGCAACTGCA 2764
```

Db 2610 AATCATTTATTCCTGAGGACAGCAGGAGGAGCAATGATCGGATATAAACCCAGTT 2669
Qy 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGGCTCTGTTTT 840
Db 2670 TTGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGGCTCTGTTTT 2729
Qy 841 CACTCATTTTCACTCTTAATTAATCATGCAACTGCA 875
Db 2730 CATGCTATTTCATCTTATAAATCTTGCAACTGCA 2764.

RESULT 16

US-09-864-761-4444
; Sequence 4444, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4444
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4

; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
US-09-864-761-4444
Query Match 53.0%; Score 704.8; DB 9; Length 1894;
Best Local Similarity 89.7%; Pred. No. 2.7e-208; Mismatches 79; Indels 9; Gaps 1;
Matches 770; Conservative 0;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTCTTAACCCGCCAAAGAGGGGGAACCTGTTTATTTT 60
Db 961 TCAAAATCGAAGAGCTTTAGACTTCTTAACCCGCCAAAGAGGGGGAACCTTTTATTTT 1020
Qy 61 AGGGGAAGAACTGCTGTAGTATGTTAAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA 120
Db 1021 AGAGGAAAAATGCTGTGTTATGTTAAATCAATCCGGAATCATCACCGAGAAAGTTAAAGA 1080
Qy 121 AATTTGAGATCGAATATATATGATAGCAGAGGAGCTTCAAAACACTGCACCTGGGCGCT 180
Db 1081 AATTCAAGGTGCAATATAACGTAAGAGAAAGAGCTGCAAAACACTGGGACCTTGGGCGCT 1140
Qy 181 CCTCAGCCAATGGATGCGCTGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 1141 CCTCAGCCAATGGATGCGCTGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1200
Qy 241 TTTACTCTCTTTGGACCTGATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 300
Db 1201 GTTACTCTCTTTGGACCTGATCTTCAACTCTCTTGTAAAGTTTGTCTTCTTCCAGAAAT 1260
Qy 301 TGAAGCTGTAAAGCTACAATATGTTTCAAAATGGAACCCAGATCAGTCCATGACTAA 360
Db 1261 CGAAGCAGTAAAACTACAATCGTTCTTCAAAATGGAGCCCGAGATCAGTCCATGAGTAA 1320
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAAATGACATTTGA 420
Db 1321 AATCTACCGAGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAAATGACATCAA 1380
Qy 421 AGTCACCCCTCCGAGGAAATCTCAACTGCAACAAACCCCTACTACATCCCAATTCAGTAGG 480
Db 1381 AGGCACCCCTCCGAGGAAATCTCAACTGCAACAAACCCCTACTACATCCCAATTCAGTAGG 1440
Qy 481 AAGCAGTTAGAGAGTTGTCAGCCACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
Db 1441 AAGCAGTTAGAGAGTTGTTGGCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGA 1500
Qy 541 GGTGAGCTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 1501 GGGGAGCTGAGAGCAGGAGTAAGTAACTAGATTTCTTAGACCACTAAGAAATCCCTAGACT 1560
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGTTGCAACTTAGCTACACCCG 660
Db 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCCGGCTTGGCAACTTAGCTACACCCCA 1620
Qy 661 ACCAATC-----AGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGGATAA 711
Db 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTTAATAGGCAAAACAGAGGATAA 1680
Qy 712 GCAATAGCCAATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATTTGGGATATAA 771
Db 1681 GAATAGCCAATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATTTGGGATATAA 1740
Qy 772 ACTCAGGCAATTCAGCCAGCAACCCCTTTGGGTCCCTCCCTCCCATTTGATGGGAG 831
Db 1741 ACCCAGGCAATTCAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGATGGGAG 1800
Qy 832 CTCGTTTTCACTCTATT 849
Db 1801 CTCGTTCTCACTCTATT 1818

RESULT 17

US-10-087-192-730/c
; Sequence 730, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 730
; LENGTH: 161334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(161334)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-730

Query Match 48.4%; Score 643.4; DB 13; Length 161334;
Best Local Similarity 85.4%; Pred. No. 4.1e-188; Indels 46; Gaps 4;
Matches 781; Conservative 0; Mismatches 88;
QY 1 TCATAATCGAAGAGCTTTAGACTTCTAACCGCCAAAGAGGGGAACTGTTTATTTT 60
DB 100898 TCATAATGAGAGCTTTAGACTTCTAACCGCCAAAGAGGGGAACTGTTTATTTT 100839
QY 61 AGGGGAAGATGCTGTAG-----TATGTTAATCA 90
DB 100838 AGGGGAAGATGCTGTATTTATTTATTTAGCGGAAGATGTTGTTATTTATTTAATCA 100779
QY 91 ATCTGGAATCATTTAGTGAAGATTTAAGAAATTTGAGATCAATATATAGAGCAGA 150
DB 100778 ATCTGGAATTTACAGAGAAAGTTGAGAAATTCGAGATTTGAATCAACGTAGAACAGA 100719
QY 151 GGACCTTC-AAAACACTGCACCTCGGGCTCTCTCAGCAATGAGATGCCCTGACTCTCC 209
DB 100718 GGAGCTTCAAAAACACCAGACCTCGGGCTCTCTCAGCCAAATGGATGCCCTGATTTCTCC 100659
QY 210 CTTCTTAGACCTCTAGCAGCTAATAATTTTACTCTCTTTTGGACCTGCTATTTCA 269
DB 100658 CTTCTTAGGATCTCTAGCAGCTTAATAATGATCTCTCTTTGGACCTGATCTTTA 100599
QY 270 ACTCTCTCTTAAGTTTGTCTCTCCAGAAATTTGAAGCTGTAAGCTCAAAATAGTTCTTC 329
DB 100598 ACTCTCTGTTAAGTTTGTCTCTCCAGAAATTTGAAGCTGTAAGCTCAAAATCGTTCTTC 100539
QY 330 AAATGGAACCCAGATGCAATGATCAATCAATCTACCGTGGACCCCTGGACCGGCTG 389
DB 100538 AAATGGAACCCAGATGCAATGATCAATCAATCTACCGTGGACCCCTGGACCGGCTTA 100479
QY 390 CTAGACTATGCTCTGATGTTATGATCAATTTGATCAATTTGATCAATTTGATCAATTTG 449
DB 100478 CTAGCCCATGCTCCAAATTTGATGATATGATCAATTTGATCAATTTGATCAATTTG 100419
QY 450 CACAACCCCTACTACACTCAATTTAGTAGGAGCAGTTAGAGCAGTTGTCAGCCAACT 509
DB 100418 CACAACCCCTACTATGCCCAATTTCCGAGGAGCAGTTAGACTGCTGTCAGCCAACT 100359
QY 510 CCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGACTGAGAGCAGACTAGTCTGA 569
DB 100358 CCCCAACAGCATTGGGTTTTCTGTTGAGTGGGGGACTGAGAGCAGGATAGTCTGA 100299

RESULT 18

US-09-854-867-385
; Sequence 385, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 385
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)...(780)
; OTHER INFORMATION: ltr17
US-09-854-867-385

Query Match 48.0%; Score 637.8; DB 10; Length 780;
Best Local Similarity 90.0%; Pred. No. 1.3e-187; Indels 19; Gaps 2;
Matches 711; Conservative 0; Mismatches 60;
QY 549 TGAGAGCAGAGCTAGCTGGATTTCTTAGGCTGACTAAGAATCCNAAAGCCTTANTCTGGGA 608
DB 1 TGAGAGCAGAGCTAGCTGGATTTCTTAGGCGCAATAGAAATCCCTAAGCTAGCTAGCTGGA 60
QY 609 AGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTACACCCGACCAATC- 667
DB 61 AGGTGACCGCTTCCACCTTTAAACACGGGCTTGCAACTTAGCTACACCCGACCAATCA 120
QY 668 -----AGAGAGCTACTAAATGCTAATCAGGCAGAAAACAGAGGTAAGCAATAGC 719
DB 121 GGTAGTAAGAGAGAGCTCACTAAATGCTAATAGGCAAAAACAGAGGTAAGAAATAGC 180
QY 720 CAATCATCTATTGCTGAGAGCAGCGGGAAGAGCAAGGATTTGGATATATAAATCAGGC 779
DB 181 CAATCATCTATCGCTGAGAGCAGCGGAGGAGCATGATCGGATATAAACCAGGC 240
QY 780 ATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATTTGAGAGCTCTGTT 839
DB 241 ATTGAGCCGCAACGGGTACCTCTTTGGGTCCCTCCCTTTGATTTGAGAGCTCTGT-- 298
QY 840 TCACCTATTTCACCTCTATTAAATCATCAACTGCACTCTTCTGCTCGGTGTTTTTTATG 899

D	b	77516	CAGAGAGCTCATTAAATGCTTAATTAGCGAAAACAGGAGGTAAAGAATAAGCAATCAT	77571
Q	y	727	CTATTCCCTTGAGAGCACAGCGGGAAAGGACAAGATTTGGGATATATAAATCAGGCATTTCAAG	786
D	b	77576	CTATTCCCTTGAGAGCACAGCGGGAGGGAACAAGCATCGGGATATAAACCCAGGCATTTCCAG	77635
Q	y	787	CGAGGACAGCAACCCCCTTTTGGGTCCCTCCCATTTGATGGAGACTCTGTTTTCACTCT	846
D	b	77636	CGGCAACGCGCAACCACTTTTGGGTCCCTCCCTTTGATGGAGACTCTG-----	77685
Q	y	847	ATTTCCTCTATTAAATCATGCAAATGCACTCTTCTGGTCCGTGTTTTTATGGCTCAAG	906
D	b	77686	CTTTCTACTCTATTAAATCTTGCACTGTCACCTCTTCTGGTCCCATGTTTCTTATGGCTCGG	77745
Q	y	907	CTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTTCGCCACCGTCACAGACCCTGCTGTAC	966
D	b	77746	CTGAGCTTTTCAGTTCGCCCTCCACCACTGCTGCTGGCCACCTTTCACAGACCACCGCTGAC	77805
Q	y	967	TTCATCCCTTTTGGATCCAGCAGAGTGTCACCTGTGCTCTGTATCCAGCAGGTTACCCAT	1026
D	b	77806	TTCCATCCCTCTGGATCTCGCAGGGTGCTGCTGTCTCTGTATCCAGCAGGCACTCAT	77865
Q	y	1027	TGGCACTCCCGATCAGGCTAAAGGCTTGCCAATTGTTCTCTGCATGGCTAAAGTGCCTGGTT	1086
D	b	77866	TGCTGCTCCCGATCGGGCTTAAGGCTTGCCAATTGTTCTCTGCACGGCTAAGTGCCTGGTT	77925
Q	y	1087	TGTCCTAATAGAACTGAAACACTGGTGACATGCGGTTTCCATGGTTCTTCCATGACCCACGG	1146
D	b	77926	CGTCTAATTAGCTGAAACACTAGTCACTGGGTTTCATGGTTCTTCCATGACCCACGG	77985
Q	y	1147	CTTCTAATGAGCTATAACACTCACCGCATGGCCCCCAAGATTCCATTCTTGGTATCTGTG	1206
D	b	77986	CTTCTAATGAGCTATAACACTCACCGCATGGCCCCCAAGATTCCATTCTTGGTATCTGTG	78037
Q	y	1207	AGGCCAAGAACCCCGAGGTGAGAGAANGTAGGCTTGCCACCAATTTGGGAAGTGGCCCACT	1266
D	b	78038	AGGCCAAGAACCCCGAGGTGAGACAACAAGAGGCTTACCACCATCTTGGAAAGTGGCCCACT	78097
Q	y	1267	GCCATTTTGTAGCGGCCCAACACATCTTTGGAGCTGTGGAGCAAGATCCCCAGTA	1326
D	b	78098	GCCATTTTGGAAAGCAGCGGCCCACTCTTTGGAGCTCTTGGAGCAAGGACACCCCACTA	78157
Q	y	1327	ACA 1329	
D	b	78158	ACA 78160	
 RESULT 20 US-10-672-764A-34/C ; Sequence 34, Application US/10672764A ; Publication No. US20040156832A1 ; GENERAL INFORMATION: ; APPLICANT: Jolly, Chris ; TITLE OF INVENTION: Immunoglobulin Compositions and Methods ; FILE REFERENCE: 13311, 1001U ; CURRENT APPLICATION NUMBER: US/10/672,764A ; CURRENT FILING DATE: 2003-09-26 ; NUMBER OF SEQ ID NOS: 68 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 34 ; LENGTH: 157090 ; TYPE: DNA ; ORGANISM: Human US-10-672-764A-34				
 Query Match 45.3%; Score 601.4; DB 18; Length 157090; Best Local Similarity 87.0%; Pred. No. 5e-175; Matches 689; Conservative 0; Mismatches 84; Indels 19; Gaps 2;				
Q	y	547	ACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCTTANCTGG	606
D	b	81195	ATTGAGAGACAGGACTAGCTGGATTTCTTAGGCGGACTAAGAAATCCCTAAGCCAGCTGG	81136


```
Qy 607 GAAGGTGACCCGATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAAT 666
Db 81135 GAAGGTGACCACTTCCACCTTTAAACACAGGGGCTTGCAACTCAGCTCACACCCGACCAAT 81076
Qy 667 C-----AGAGAGCTCCTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATA 717
Db 81075 CAGATAGTAAGAGAGCTCCTAAATGCTAATTAGGCAAAAGCAGGGTAAAGAAATA 81016
Qy 718 GCCAATCATCTATTGCTGAGAGACACAGCGGAAAGGACAAGGATTTGGGATATAAATCAG 777
Db 81015 GCCAATCATCTATCACCTGAGAGACACAGCGGAGGGAACAATGATCGGGATATAAACCAG 80956
Qy 778 GCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGT 837
Db 80955 GCATTGAGCGGGTAACAGCTACCCCTTTTGGGTCCCTCCCTTTGATGGGAGCTCTGT 80896
Qy 838 TTTCACCTCTATTCTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTA 897
Db 80895 -----CTTCACCTCTATTCTCTGCAACTGCACTCTTCTGGTCCATGTTGTA 80846
Qy 898 TGGCTCAAGCTGAGCTTTTGGTTCGGCATCCACCACTGCTGTTTGGCACCCGTCAGACCC 957
Db 80845 CGGCTCGAGCTTCGCTTTTCGCTCAACCGTCCACCACTGCTGTTTGGCCCCGTTGGAGACCC 80786
Qy 958 GCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTCATCCAGCA 1017
Db 80785 ACCGCTGGCTTCCATCCCTTCGGATCCATCAGGGTATCCGCTGTCTCTGATCCAGCA 80726
Qy 1018 GGTACCACTTGGCACTCCCGATCAGGCTAAAGGCTTGGCAATGTTCTTCGATGCTAAGT 1077
Db 80725 GGGCCCACTTGGCACTCTCGATCGGGCAAAATGCTCGCAATGTTCTTCGACGCGTAAGT 80666
Qy 1078 GCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCATGGTCTCTTCCAT 1137
Db 80665 GCCTGGGTTCTGCTCTAAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCTCTCCGT 80606
Qy 1138 GACCCACGGCTTCTAATAGAGCTAACAACACTCACCGCATGGCCCAAGATTCATTCCCTG 1197
Db 80605 GACCCACGGCTTCTAATAGAGCTAACAACACTCACTGATGAGCCCAAGCTTCCATTCCTTG 80546
Qy 1198 GTATCTGTGAGCCCAAGAACCCAGGTCAGAGAANGTAGGCTTGCCACCAATTTGGGAAG 1257
Db 80545 GAATCCGTGAGGCCAAGAACCCAGGTCAGAGAACATGAGGCTTGCCACCACTTTGGAAG 80486
Qy 1258 TGGCCCACTGCATTTTGGTAGCGGCCACCAACCATCTTGGGAGCTGCGGAGCAAGGAT 1317
Db 80485 TGGCTGCGGTCATTTTGGAAAGCGGCTTGCCACCACTTTGGGAGCTCTGGGAGCAAGGC 80426
Qy 1318 CCCCCAGTAACA 1329
Db 80425 CCCCCAATAACA 80414
```

```
US-10-003-806-6
Query Match 43.6%; Score 579.8; DB 13; Length 180557;
Best Local Similarity 86.0%; Pred. No. 2.9e-168;
Matches 683; Conservative 0; Mismatches 90; Indels 21; Gaps 3;

Qy 547 ACTGAGAGACAGGACTAGCTGGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCCTANCTGG 606
Db 58991 AGTGAGAGACAGGACTAGCTGGGATTTCTTAGGCGGACTAAGAAATCCCTNAAAGCTAGCTGG 58932
Qy 607 GAAGGTGACCCGATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAAT 666
Db 58931 GAAGGTGACCCGCTTCCACCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAAT 58872
Qy 667 C-----AGAGAGCTCCTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATA 717
Db 58871 CAGATAGTAAGAGAGCTCCTAAATGCTAATTAGGCAAAACAGGAGGTAAAGAAATA 58812
Qy 718 GCCAATCATCTATTGCTGAGAGACACAGCGGAAAGGACAAGGATTTGGGATATAAATCAG 777
Db 58811 GCCAATCATCTATTGCTGAGAGACACAGGAGGGAACAATGATCAGCATATAAACCCAG 58752
Qy 778 GCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGT 837
Db 58751 GCATCTGAGCCAGCAACAGCTACGCTCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGT 58692
Qy 838 TTTCACCTCTATTCTACCTCTAATTAATCATCAACTGCA--CTCTTCTGGTCCGTGTTTTT 895
Db 58691 -----CTTCACCTCTAATTAATCTTGAGCTGCACTCTCTTTTGGTCTACATTTGT 58642
Qy 896 TATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGAC 955
Db 58641 CATGGTTCGAGCTGAGCTTTCTCTCGCGCTCCACCACTGCTGTTTGGCGCTGTCGACAG 58582
Qy 956 CCGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGC 1015
Db 58581 CTGCTGCTGACTTCCATCCGTCAGATCCGCAAGGCTGTGCTGTGCTCTGATCCAGC 58522
Qy 1016 GAGGTACCCATGCGCACTCCCGATCAGGCTAAAGGCTTGCCATTTGTTCTGCATGGCTAA 1075
Db 58521 GAGCGGCCAATTAAGCTTCCAGCTCGGCTTAAGGCTTGCCACTGTTCCACACGGCTAA 58462
Qy 1076 GTGCTCGGGTTGTCCTAATAGAACTGAACACTGGTCACTGGGTTCCATGTTCTCTTCC 1135
Db 58461 GTGTCGCGGTTGCTCTTAATCAAGCTGAACACTAGTCACTGGGTTCCACGGTCTCTTCC 58402
Qy 1136 ATGACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCT 1195
Db 58401 ATGACCCACGGCTTCTAATAGAGCTATAAACAACACTCACTGATGGCCCAAGATTCCATTCCT 58342
Qy 1196 TGGTATCTGTGAGGCCAAGAACCCAGGTCAGAGAANGTAGGCTTGCCACCATTTGGGA 1255
Db 58341 TGGAATCCGTGAGGCCAAGAACCCAGGTCAGAGAACAGGAGGCTTGCCACCATCTTGA 58282
Qy 1256 AGTGCCCACTGCGCATTTTGGTAGCGGCCCAACCACTATCTTTGGGAGCTGTGGGAGCAAG 1315
Db 58281 AGCAGCCCGCGCATTTTGGAAAGTGCGCCACCACTCTTTGGGAGCTCTGGGAGCAAG 58222
Qy 1316 ATCCCCCAGTAACA 1329
Db 58221 ACCACCTGGTAACA 58208
```

```
RESULT 22
US-10-003-806-9/c
; Sequence 9, Application US/10003806
; Publication No. US20020119929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
```

```
US-10-003-806-6/c
; Sequence 6, Application US/10003806
; Publication No. US20020119929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,872
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 180557
; TYPE: DNA
; ORGANISM: Human
```



```
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,872
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 180557
; TYPE: DNA
; ORGANISM: Human
US-10-003-806-9

Query Match      43.6%; Score 579.8; DB 13; Length 180557;
Best Local Similarity 86.0%; Pred. No. 2.9e-168;
Matches 683; Conservative 0; Mismatches 90; Indels 21; Gaps 3;

QY 547 ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGG 606
DB 58991 ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCGGACTAAGAAATCCCTAAGCTAGCTGG 58932

QY 607 GAAGGTGACCGATCCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCGGACCAAT 666
DB 58931 GAAGGTGACCGTTCCACCTTTAAACATGGGGTTGCAACTTAGCTCACACCGGACCAAT 58872

QY 667 C-----AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATA 717
DB 58871 CAGATAGTAAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATA 58812

QY 718 GCGAATCATCTATTGCTGAGAGCACAGCGGGAAGGACAAGGATTTGGGATATAAACTCAG 777
DB 58811 GCGAATCATCTATTGCTGAGAGCACAGCAGGAGGAGCAATGATCAGCATATAAAACCCAG 58752

QY 778 GCATTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGT 837
DB 58751 GCATCTGAGCCAGCAACAGCTACGCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTGT 58692

QY 838 TTTCACCTCTATTTCACCTCTATTAAATCATGCACTGCA--CTCTCTGCTCGGTGTTTT 895
DB 58691 -----CTTCACCTCTATTAAATCTTGCAGCTGCACCTCTCTTTTGGTCTACATTTGT 58642

QY 896 TATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAAGAC 955
DB 58641 CATGGTTCGAGCTGAGCTTCTCTCGCGCTCCACCACTGCTGTTTGGCGCTGTCGAGAC 58582

QY 956 CGCTGCTGACTTCATCCCTTTGGATCCAGAGAGTGCCACTGTGCTCTGATCCAGC 1015
DB 58581 CTGCTGCTGACTTCATCCGTCAGATCCGCGCAAGGTGCTGTGCTGCTGATCCAGC 58522

QY 1016 GAGGTACCAATGGCCACTCCGATCAGGCTAAAGGCTTGCCATTGTTCTGCAATGGCTAA 1075
DB 58521 GAGGCGCCCATTAACAGCTCCGAGTCGGGCTAAAGGCTTGGCACTGTTCCACAGGCTAA 58462

QY 1076 GTGCTCGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCATGGTTCTCTTCC 1135
DB 58461 GTGTCGGGTTGTCTTAATCAAGCTGAACACTAGTCACTGGGTTCACAGGTTCTCTTCC 58402

QY 1136 ATGACCCAGCGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCT 1195
DB 58401 ATGACCCAGCGCTTCTTAATAGAGCTATAACACTCACCTGATGGCCCAAGATTCATTCCT 58342

QY 1196 TGGTATCTGTGAGGCCAAGAACCCAGGTGAGGAGGTGAGGCTTGCCACCAATTTGGGA 1255
DB 58341 TGGAAATCCGTGAGGCCAAGAACCCAGGTGAGGAGGTGAGGCTTGCCACCAATTTGGGA 58282

QY 1256 AGTGGGCCACTGCGCATTTTGGTAGCGGCCCAACCATCTTTGGGAGGTGTTGGGAGCAAGG 1315
DB 58281 AGCAGCCCGCGCCATTTTGGAGTGGGCCCAACCATCTTTGGGAGGTGTTGGGAGCAAGG 58222

QY 1316 ATCCCCCAGTAACA 1329
DB 58221 ACCACCTGGTAACA 58208
```

RESULT 23

```
US-10-017-117-1
; Sequence 1, Application US/10017117
; Publication No. US20030124535A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MMI-008
; CURRENT APPLICATION NUMBER: US/10/017,117
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: No. US20030124535A1 yet assigned
; PRIOR FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 161671
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(161671)
; OTHER INFORMATION: n = a,t,c or g
US-10-017-117-1

Query Match      42.9%; Score 570.2; DB 15; Length 161671;
Best Local Similarity 88.5%; Pred. No. 2.7e-165;
Matches 655; Conservative 0; Mismatches 71; Indels 14; Gaps 3;

QY 547 ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGG 606
DB 74117 ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCGGACTAAGAAATCCCTAAGCTAGCTGG 74176

QY 607 GAAGGTGACCGATCCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCGGACCAAT 666
DB 74177 GAAGGTGACCGATCCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCGGACCAAT 74236

QY 667 CAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCAT 726
DB 74237 CAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCAATCAT 74296

QY 727 CTATTGCTGAGAGACAGCGGAGGACAGGATTTGGGATATAAACTCAGGCAATTCAG 786
DB 74297 CTATTGCTGAGAGACAGCGGAGGACAGGATTTGGGATATAAACTCAGGCAATTCAG 74356

QY 787 CCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTTCACCTCT 846
DB 74357 CCGGCAATGGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGT-----74407

QY 847 ATTTCACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTTATGGCTCAAG 906
DB 74408 -TTTCACTCTATTAAATCTTGCACACTGCACTCTTCTGGTCCATGTTTGTACGGCTGGAG 74466

QY 907 CTGAGCTTTTGTGGCCATCCACCACTGCTG--TTTGCCACCGTTCAGAGCCCGCTGCTGA 965
DB 74467 CTGAGCTTTTGTGGCCATCCACCACTGCTGTTTTCGGCCCGTGGCAGACTCGCCACTGA 74526

QY 966 CTTTCCATCCCTTTGGATCCAGGAGGTGTCCTGCTGCTCTGATCCAGGAGGTACCCCA 1025
DB 74527 CTTTCCATCTTCGCGATCCGGAGGGGTGTCCTGCTGCTGCTGATCCAGGAGGTACCCCA 74586

QY 1026 TTGCCACTCCCGATCAGGCTAA---AGGCTTGCCATTGTTCTGCAATGGCTAAGTGCCTG 1082
DB 74587 CTGCCATTCGGATCGGGCTAAATTAGGCTTGCCATTGTTCTGCTGCTGCTGCTGCTGCTG 74646

QY 1083 GGTTCCTCTAATAGAACTGAACACTGAGTTCATGGGTTCCATGGTTCCTTCCATGACCC 1142
DB 74647 GGTTCCTCTAATCGAGCTGAACACTAGTCACTGGGTTCCACGGTTCCTTCTCGTGAACC 74706

QY 1143 ACGGCTTCTAATAGAGCTATAACACTCAGCGATGCCCAAGATTTCATTTCTTGGTATC 1202
DB 74707 ACGGCTTCTAATAGAGCTATAACACTCAGCGATGCCCAAGATTTCATTTCTTGGTATC 74766

QY 1203 TGTGAGGCCAAGAACCCCGAGGTGAGAGAAAGTGGAGCTTTGCCACCATTTGGGAAGTGGCC 1262
```

Db 74767 CGTGAGGCCAAGAACCCCGAGGTCAGAGAACGAGGCTTGTACCATCTTTGGAAGCAGTC 74826
Qy 1263 CACTGCCATTTTGGTAGCGG 1282
Db 74827 CACCGCCATCTTGAAGCTG 74846
RESULT 24
US-10-240-425-1102
; Sequence 1102, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzell, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1102
; LENGTH: 134292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AL031983
US-10-240-425-1102
Query Match 41.0%; Score 545.4; DB 17; Length 134292;
Best Local Similarity 85.9%; Pred. No. 1.3e-157;
Matches 680; Conservative 0; Mismatches 89; Indels 23; Gaps 6;
Qy 547 ACTGAGACAGCAGCTAGCTGATTTCTTAGCTGACTAAGAAATCCNNAAGCTTACTGG 606
Db 54610 AATGAGACAGCAGCTAGCTGATTTCTTAGCCGCACTAAGAAATCCCTAAGCTAGCTGG 54669
Qy 607 GAAGGTGACCGCATCATCTTTAAACATGGGCTTGCACCTTAGCTCACCCGACCAAT 666
Db 54670 GGAGGTGACTGCATCCACCTTTAAACACGGGCTTGCACCTTAGCTCACACCTGACCAAT 54729
Qy 667 C-----AGAGGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATA 717
Db 54730 CAGGTAGTAAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAGATATA 54789
Qy 718 GCMAATCATCTATTGCTCAGAGCAGCGGAAGACAAAGATTGGGATATAAACTCAG 777
Db 54790 GCAATCATCTATTGCTGAGATCACAGCGGAGGACAAATGATCGGATATAACACAG 54849
Qy 778 GCATTCGAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGT 837
Db 54850 GCATTCGAGCCAGCAAC-GCTACCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTGT 54908
Qy 838 TTTCACCTCTATTTCATCTATTAAATCATGCACTGCACTCTCTGGTCCGTTTCTTTA 897
Db 54909 -----CTTCACCTACTAAATCTTGGCAACTGCACTCTCTGGTCTATGTTTCTTA 54958
Qy 898 TGCTCAAGCTGAGCTTTTGTTCGCCATCCACCATGCTGTTTGGCCACCGTCCAGACCC 957
Db 54959 CGGCTCGAGGTGAGCTTTGCTTGGCATCCACCATGCGTTCGCCACCGTCCGACCC 55018
Qy 958 GCTGCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCACTGTGCTCTCTGATCCAGCA 1017
Db 55019 GCGGCTGACTTCATCCCTC-GGATCTGGCAGGGTGTCCGCTGTGTCTTCTGAACCAAGTGA 55077

Qy 1018 GGTACCATTTGCCACTCCCGATCAGGCTAAAGGCTTGGCAATTGTTCTCTGCATGGCTAAGT 1077
Db 55078 GGCGCCCATTTGCCGCTCCTGATTGGGCTAAAGCGGTACCATTTGTT-CTGCACGGCTAAGT 55136
Qy 1078 GCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCTCACTGGGTTCCATGGTTCCTTCCAT 1137
Db 55137 GCCCAGGTTCTTCTTAATCGAGCTGAACACTAGTCACTGGG-TCCACAGTTCTCTTCCGT 55195
Qy 1138 GACCCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCCATTCCTTG 1197
Db 55196 GACCCATGGCTTCTTAATAGAGCTGTAAACACCAACCATGACCCCAAGATTCCCATTCCTTG 55255
Qy 1198 GTATCTGTGAGGCCAAGAACCCCGAGGTCAGAGAANGTGAGGCTTGCCACCATTTGGGAAG 1257
Db 55256 GAATCCATGAGGCCAAGAACCCCGAGGTCAGAGAACACGAGGCTTGCCACCATCTTGGGAAG 55315
Qy 1258 TGGCCCACTGCCATTTTGTAGCGGCCCAACCACTCTTGGGAGCTGTGGAGCAAGAT 1317
Db 55316 CGGCTGCGGCCATTTTGGGAAGCAGCCACCATCTTGGAGAGCTCTGGGAGCAAGAC 55375
Qy 1318 CCCCCAGTAACA 1329
Db 55376 CCCCCTGGTAACA 55387
RESULT 25
US-08-979-847-102
; Sequence 102, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

Qy	797	CAACCCCTTTGGGCCCTCCCATGATGATGGAGACTCTGTTTCACTCTATTTCACCTCT	856
Db	541	CAACCCCTTTGGGCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTATTTCACCTCT	600
Qy	857	ATTAAATCATGCAACTGCA	875
Db	601	ATTAAATCTTGCACCTGAA	619

```

RESULT 27
US-120-120-15
; Sequence 15, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, SCOTT R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKDREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR D
; FILE REFERENCE: PT-113 PCT
; CURRENT APPLICATION NUMBER: US/10/272-2
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,771
; 60/184,693; 60/184,771; 60/184,818; 60/184,834; 60/184,876; 60/184,883; 60/184,888; 60/184,891; 60/184,894; 60/184,897; 60/184,900; 60/184,903; 60/184,906; 60/184,909; 60/184,912; 60/184,915; 60/184,918; 60/184,921; 60/184,924; 60/184,927; 60/184,930; 60/184,933; 60/184,936; 60/184,939; 60/184,942; 60/184,945; 60/184,948; 60/184,951; 60/184,954; 60/184,957; 60/184,960; 60/184,963; 60/184,966; 60/184,969; 60/184,972; 60/184,975; 60/184,978; 60/184,981; 60/184,984; 60/184,987; 60/184,990; 60/184,993; 60/184,996; 60/184,999; 60/185,002; 60/185,005; 60/185,008; 60/185,011; 60/185,014; 60/185,017; 60/185,020; 60/185,023; 60/185,026; 60/185,029; 60/185,032; 60/185,035; 60/185,038; 60/185,041; 60/185,044; 60/185,047; 60/185,050; 60/185,053; 60/185,056; 60/185,059; 60/185,062; 60/185,065; 60/185,068; 60/185,071; 60/185,074; 60/185,077; 60/185,080; 60/185,083; 60/185,086; 60/185,089; 60/185,092; 60/185,095; 60/185,098; 60/185,101; 60/185,104; 60/185,107; 60/185,110; 60/185,113; 60/185,116; 60/185,119; 60/185,122; 60/185,125; 60/185,128; 60/185,131; 60/185,134; 60/185,137; 60/185,140; 60/185,143; 60/185,146; 60/185,149; 60/185,152; 60/185,155; 60/185,158; 60/185,161; 60/185,164; 60/185,167; 60/185,170; 60/185,173; 60/185,176; 60/185,179; 60/185,182; 60/185,185; 60/185,188; 60/185,191; 60/185,194; 60/185,197; 60/185,200; 60/185,203; 60/185,206; 60/185,209; 60/185,212; 60/185,215; 60/185,218; 60/185,221; 60/185,224; 60/185,227; 60/185,230; 60/185,233; 60/185,236; 60/185,239; 60/185,242; 60/185,245; 60/185,248; 60/185,251; 60/185,254; 60/185,257; 60/185,260; 60/185,263; 60/185,266; 60/185,269; 60/185,272; 60/185,275; 60/185,278; 60/185,281; 60/185,284; 60/185,287; 60/185,290; 60/185,293; 60/185,296; 60/185,299; 60/185,302; 60/185,305; 60/185,308; 60/185,311; 60/185,314; 60/185,317; 60/185,320; 60/185,323; 60/185,326; 60/185,329; 60/185,332; 60/185,335; 60/185,338; 60/185,341; 60/185,344; 60/185,347; 60/185,350; 60/185,353; 60/185,356; 60/185,359; 60/185,362; 60/185,365; 60/185,368; 60/185,371; 60/185,374; 60/185,377; 60/185,380; 60/185,383; 60/185,386; 60/185,389; 60/185,392; 60/185,395; 60/185,398; 60/185,401; 60/185,404; 60/185,407; 60/185,410; 60/185,413; 60/185,416; 60/185,419; 60/185,422; 60/185,425; 60/185,428; 60/185,431; 60/185,434; 60/185,437; 60/185,440; 60/185,443; 60/185,446; 60/185,449; 60/185,452; 60/185,455; 60/185,458; 60/185,461; 60/185,464; 60/185,467; 60/185,470; 60/185,473; 60/185,476; 60/185,479; 60/185,482; 60/185,485; 60/185,488; 60/185,491; 60/185,494; 60/185,497; 60/185,500; 60/185,503; 60/185,506; 60/185,509; 60/185,512; 60/185,515; 60/185,518; 60/185,521; 60/185,524; 60/185,527; 60/185,530; 60/185,533; 60/185,536; 60/185,539; 60/185,542; 60/185,545; 60/185,548; 60/185,551; 60/185,554; 60/185,557; 60/185,560; 60/185,563; 60/185,566; 60/185,569; 60/185,572; 60/185,575; 60/185,578; 60/185,581; 60/185,584; 60/185,587; 60/185,590; 60/185,593; 60/185,596; 60/185,599; 60/185,602; 60/185,605; 60/185,608; 60/185,611; 60/185,614; 60/185,617; 60/185,620; 60/185,623; 60/185,626; 60/185,629; 60/185,632; 60/185,635; 60/185,638; 60/185,641; 60/185,644; 60/185,647; 60/185,650; 60/185,653; 60/185,656; 60/185,659; 60/185,662; 60/185,665; 60/185,668; 60/185,671; 60/185,674; 60/185,677; 60/185,680; 60/185,683; 60/185,686; 60/185,689; 60/185,692; 60/185,695; 60/185,698; 60/185,701; 60/185,704; 60/185,707; 60/185,710; 60/185,713; 60/185,716; 60/185,719; 60/185,722; 60/185,725; 60/185,728; 60/185,731; 60/185,734; 60/185,737; 60/185,740; 60/185,743; 60/185,746; 60/1
```

```

;
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:429446.1:2000FEB18
; US-10-220-120-15

Query Match 40.3%; Score 536.2; DB 17; Length 849;
Best Local Similarity 81.4%; Pred. No. 6.5e-156;
Matches 650; Conservative 0; Mismatches 135; Indels 14; Gaps 2;

Qy 208 CCCCTCTTAGGACCTCTAGCAGCTATAATATTTTACTCCTCTTTTGAGCCCTGATCTT 267
Db 46 CACCATCTTTGGGAGCTCTAAGAACAAAGACCCACCCTAAACATTTGTGGCACTGTACG 105
Qy 268 CAATCTTCCTTTAAAGTTTCTCTCTCCAGAAATTGAAGCTGTAAAGCTACAAATAGTTCT 327
Db 106 GGAATTCCTCAAAGTGTGTCTCTCCAGAACTCAAACCTGTAACAACTACAAATTTGTTCT 165
Qy 328 TCAAAATGGAAACCCAGATGCAAGTCATGACTAAATCTACGGTGGACCCCTGGACCGGCC 387
Db 166 TCAAAATGGAGCCAGATGCAAGTCATGACTAAAGATCTACACAGACCCCTTGGACCGGCC 225
Qy 388 TGTCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAGGAAATCTCAAC 447
Db 226 TGTCTAGCCCATGCTCCGATGTTGATGACATCGAAGGCGCCCTCCCAAGGAAATCTCAAC 285
Qy 448 TGCACAAACCCCTACTACACTCCAATTCAGTAGGAAGCAGTTAGAGCAGTTGTCAGCCCAAC 507
Db 286 TGCATGACCCCTACTACGCCCCAATTCGCGAGGAAGCAGTTAGAGCAGTCAATGGCCCAAC 345
Qy 508 CTCCCAACAGTACTTGGGTTTTCCTCTTGAGAGGGTGGACTGAGAGACAGACTAGCTG 567
Db 346 CTCCCAACAGCACTTGAGTTTTCCTGTGTAGAGGGAGAACTGAGAGACAGCACTAGCTG 405
Qy 568 GATTTCTCTAGCTGACTAAGAATCCNNAAGCCTANCTGGGAAGGTGACCGCATCATCTT 627
Db 406 GATTTCCAGGCCCAACTAAGAACTCCCTAAGCCTATCTGGGAAGGTGACCGTATCCACCTT 465
Qy 628 TAAACATGGGGCTTGCAACTTTAGCTCACACCCGACCAATCA-----GAGAGCTCAC 678
Db 466 TAAACATAGGGCTTGCAACTTTAGCTCACACCCGACTAATCAGATAGTAAGGAGAGCTCAC 525
Qy 679 TAAATCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCTATTTGCTCTGAG 738
Db 526 TAAATGCTTAATTAGGCNAAACAGGAGGTAAAGAAATAGCAATCATCTATTGCTCTGAG 585
Qy 739 AGCACAGCGGAAGGAACAAGATTTGGGATATAAACTCAGGCATTTCAAGCCGACACACGA 798
Db 586 AGCACAGCAGGAGGACAATGATCAGGATATAAAACCCAGGCAATTCAGCCGACCAATGGCT 645
Qy 799 ACCCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTTTCACTCTATTTCACTCTAT 858
Db 646 ACCCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTTCATGCTATTTTAAAGTCT 705
Qy 859 TAAATCATGCAACTGCACTCTTCGTCCGTGTTTTTTATGGCTCAAGCTGAGCTTTTGT 918
Db 706 TGCAC-----TGCACACTCTTCTGATCCATGTTGTTTACAGCTTTGAGCTGAGCTTTTGC 760
Qy 919 TCGGCATCCACACTGCTGTTTGGCACCGTCACAGACCGGCTGCTGACTTCCATCCCTTT 978
Db 761 TCACCATCCACAAATGCTCTTTGGCACCATCACAGACCCGCACTGCTTCCATCCCTCT 820
Qy 979 GGATCCACGAGGTGTCCA 997
Db 821 GGATCCACGAGGGTTTCA 839

RESULT 28
US-10-719-993-6792/c
; Sequence 6792, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

```

US-10-719-993-6792

Query Match	38.7%;	Score 513.8;	DB 19;	Length 80032;
Best Local Similarity	83.0%;	Pred. No. 7.2e-148;		
Matches 659;	Conservative	3;	Mismatches 108;	Indels 24; Gaps 6;
Qy	549	TGAGAGACAGGACTAGCTGGATTTCCTTAGGCTGACTAAGAATCCNAAAGCCTANCT--GGG	607	
Db	31830	TGAGAGACAGGACTAGCTGGATTTCCTTAGGCGGACTAAGATTCCTAAGCCTAGCTGGG	31771	
Qy	608	AAGGTGACCGCATCCATCTTTAAACATCGGGCTTGGCAACTTGTAGCTCACACCCGACCAATC	667	
Db	31770	AAGGTGACTGCACCTTACCTTTAAAVATGGGCTGTTACTCAGCTCACACCGGCCAATC	31711	
Qy	668	-----AGAGAGCTCACTAAATGTCTAATCAGGC--AAAAACAGGAGGTAAAGCAATA	717	
Db	31710	AGGTCATAAAGAGAGCTCACTAAAAATACCAATAGGCTTAAAGCAGGAGGTAAAGAAATA	31651	
Qy	718	GCCAACTCATTTATGCTTGAGAGACACGCGGAGGACAGGATTTGGGATATAAACTCAG	777	
Db	31650	GTCAATCATCTTATCACTTGAGAGACACGCGGAGGAGCAAGATCCAGATATAAACCCAG	31591	
Qy	778	GCATTCAGCCGACGACGAAACCCCTTTGGGTCCCCTCCCATTTGTATGGAGCTCTGT	837	
Db	31590	GCATTTGAGCGGCACTGCGCAACCTGCTTTGGGTCCGCTCCGTTGTATGAGAGTTCTGT	31531	
Qy	838	TTTCACTCTATTTCACCTCTATTAAATCATGCACTGC--ACTCTCTGGTCCGCTGTTTTT	895	
Db	31530	-----TTTCACTCTATTAAATCTTGGCAACTTGCATCTCTCTGGTCTGTGTTTTAT	31481	
Qy	896	TATGGCTCAAGCTGAGCTTTTCTTCGCCATCCACCACTGCTTTTGGCCACCGTCACAGAC	955	
Db	31480	TCGTGGCTCAAGCTGAGCTTTTCTTCACCGTCCACCACTGCTGATCTCCGCCATCGCAGAC	31421	
Qy	956	CGCTGCTGACTTCCATCCCTTTTGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGC	1015	
Db	31420	CGCTGCTGACTTCCACCCTCCAGATCCAGCAGGCTGTCTGCTGTCTGATCCAGC	31361	
Qy	1016	GAGGTACCATTTGCCACTCCCGATCAGCTAAAGGCTTTGCCATTTCTTCTCATGGCTAA	1075	
Db	31360	AAGGTCCCATTTGCGCTCCCAATCGGGCTAGAGGCTCGCCATTTGTTCTGATGGCTAA	31301	
Qy	1076	GTCCCTGGGTTTCTCCTAATAGAACTGAACACTGGTCTACTGGGTTCCATGGTTCTCTTCC	1135	
Db	31300	GTGCCCAGGTTCTGCTCTAATCCAGCTGAACACTAGTCGCTGGGTTCCACGGTTCTCTTCC	31241	
Qy	1136	ATGACCCACGGCTTCTTAATAGAGCTATAACACTCAACCGCATTGGCCCAAGATTCATTCCT	1195	
Db	31240	ATGACCCATGGCTTCTTAATAGAGCTATAACACTCAACCGCATTGGCCCAAGGTTCCATTCCT	31181	
Qy	1196	TGTTATCTGTGAGGCCAAGAACCCAGGTCAGAGAANGTGAGGCTTGGCCACCATTTGGGA	1255	
Db	31180	TGAAATCCGTGAGGCCAAGAACCCCAAGATCAGAGAAACAGAGGCTTCCACCATTTGGG	31121	
Qy	1256	AGTGGGCCACTGTCATTTTGGTAGGGGCCCAACCATCTTTGGGAGCTGTGGGAGCAAGG	1315	
Db	31120	AGTGGCCGCCACCATCTAGGTAGC-ACCTGGCCACTATCTTTGGAAGCTCTAAGAACAAAG	31062	
Qy	1316	ATCCCCCAGTAAACA	1329	

Db	31180	TGGAATCCGTGAGGCCCAAGAAACCCACAGATCAGAGAACAACAAGGCTTTCACACCATCTTGGG	31182
Qy	1256	AGTGGGCCACTGCGCATTTTGGTAGGGGCCACACCCATCTTTGGGAGCTGTGGGAGCAAGG	1315
Db	31120	AGTGTCCGCGCACCATCTAGGTAGC-ACCTGCCACTATCTTTGGAAGCTCTAAGAACAAAG	31062
Qy	1316	ATCCCCCAGTAACA	1329

Db 175341 GTATGGAGCTCTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCA 175391
|||||
RESULT 34
US-10-741-600-17995
; Sequence 17995, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17995
; LENGTH: 283351
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)...(283351)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-10)
US-10-741-600-17995
Query Match 36.0%; Score 478.6; DB 19; Length 283351;
Best Local Similarity 90.1%; Pred. No. 1.3e-136;
Matches 534; Conservative 1; Mismatches 52; Indels 6; Gaps 2;
Qy 283 GTTTGCTCTTCCAGAAATTGAAGCTGTAAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGACCCCA 342
Db 174805 GTTTGCTCTTCCAGAAATTGAAGCTGTAAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGACCCCA 174864
Qy 343 GATGAGTCCATGACTTAAATCTACCGTGGACCCCTGGACCGGCTGCTAGCATATGCTC 402
Db 174865 GATGAGTCCATGACTTAAATCTACCGTGGACCCCTGGACCGGCTGCTAGCATATGCTC 174924
Qy 403 TGATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAAATCTCAACTGACAAACCCCTACT 462
Db 174925 CAATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAAATCTCAACTGACAAACCCCTACT 174984
Qy 463 ACATCTCAATTCAGTAGGACAGTTAGACGAGTTGTACGCCAACCCTCCCAACAGTACT 522
Db 174985 ATGCCCAATTCAGTAGGACAGTTGTAGACGAGTTGTAGCGAGTCAITGGCCCAACCCTCCCAACAGC 175044
Qy 523 TGGGTTTTCTCTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGA 582
Db 175045 TGGGTTTTCTCTTGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCGCA 175104
Qy 583 CTAAGAATCCNNAAGCTTANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTG 642
Db 175105 CTAAGAATCCCTAAGCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 175164
Qy 643 CAACTTAGCTCACACCGGACCAATCAGAGAGCTCACTAAATGCTAAATCAGSCAAAACA 702
Db 175165 CAACTTAGCTCACAC---CCAAACAGAGAGTTCACTAAATGCTAAATGAGCAAAATA 175220
Qy 703 GGAGGTAAAGCAATAGCCCAATCATCTATTGCTGAGAGACAGCGGGAGGACAGGAT 762
Db 175221 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGAGAGACAGCGGGAGGACAGGATC 175280
Qy 763 GGGATATAAATCTCAGGCAATTCAGAGAGTTCAGAGTGGCAACCGCCCTTTGGGTCCCTCCCAT 822
Db 175281 AGGATATAAACCCAGGCAATTCAGAGTGGCAACCGCCCTTTGGGTCCCTCCCTTT 175340
Qy 823 GTATGGAGCTCTGTTTTCACTCTATTTCACCTCTATTAAATCATGCAACTGCA 875
Db 175341 GTATGGAGCTCTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCA 175391
|||||
RESULT 35

US-10-719-993-6815
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719.993
; CURRENT FILING DATE: 2003-11-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-10)
US-10-719-993-6815
Query Match 36.0%; Score 478.6; DB 18; Length 1980090;
Best Local Similarity 90.1%; Pred. No. 3.8e-136;
Matches 534; Conservative 1; Mismatches 52; Indels 6; Gaps 2;
Qy 283 GTTTGCTCTTCCAGAAATTGAAGCTGTAAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGACCCCA 342
Db 177083 GTTTGCTCTTCCAGAAATTGAAGCTGTAAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGACCCCA 177142
Qy 343 GATGAGTCCATGACTTAAATCTACCGTGGACCCCTGGACCGGCTGCTAGCATATGCTC 402
Db 177143 GATGAGTCCATGACTTAAATCTACCGTGGACCCCTGGACCGGCTGCTAGCATATGCTC 177202
Qy 403 TGATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAAATCTCAACTGACAAACCCCTACT 462
Db 177203 CAATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAAATCTCAACTGACAAACCCCTACT 177262
Qy 463 ACATCTCAATTCAGTAGGACAGTTAGAGCAGTTGTACGCCAACCCTCCCAACAGTACT 522
Db 177263 ATGCCCAATTCAGTAGGACAGTTGTAGAGCAGTTGTAGGCCAACCCTCCCAACAGC 177322
Qy 523 TGGGTTTTCTCTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGA 582
Db 177323 TGGGTTTTCTCTTGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCGCA 177382
Qy 583 CTAAGAATCCNNAAGCTTANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTG 642
Db 177383 CTAAGAATCCCTAAGCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 177442
Qy 643 CAACTTAGCTCACACCGGACCAATCAGAGAGCTCACTAAATGCTAAATCAGGCAAAAACA 702
Db 177443 CAACTTAGCTCACAC---CCAAACAGAGATTCACATAAATGCTAAATGAGCAAAATA 177498
Qy 703 GGAGGTAAAGCAATAGCCCAATCATCTATTGCTGAGAGACAGCGGGAGGACAGGAT 762
Db 177499 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGAGAGACAGCGGGAGGACAGGATC 177558
Qy 763 GGGATATAAATCTCAGGCAATTCAGAGAGTTCAGAGTGGCAACCGCCCTTTGGGTCCCTCCCAT 822
Db 177559 AGGATATAAACCCAGGCAATTCAGAGTGGCAACCGCCCTTTGGGTCCCTCCCTTT 177618
Qy 823 GTATGGAGCTCTGTTTTCACTCTATTTCACCTCTATTAAATCATGCAACTGCA 875
Db 177619 GTATGGAGCTCTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCA 177669
|||||
RESULT 36
US-10-741-600-17676
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.


```

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17676
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

Query Match      36.0%; Score 478.6; DB 19; Length 1980090;
Best Local Similarity 90.1%; Pred. No. 3.8e-136;
Matches 534; Conservative 1; Mismatches 52; Indels 6; Gaps 2;

QY 283 GTTTGTCTCTCCAGAAATTGAAGCTGTAAAGCTACAAATAGTCTTCAAAATGGAACCCCA 342
Db 1777083 GTTTGTCTCTCCAGAAATTGAAGCTGTAAAGCTACAAATAGTCTTCAAAATGGAACCCCA 1777142

QY 343 GATCGAGTCCATGACTTAAATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTC 402
Db 1777143 GATCGAGTCCATGACTTAAATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTC 1777202

QY 403 TGATGTTAATGACTTGAAGTCAACCTCCCGAGGAAATCTCAACTGCAACACCCCTACT 462
Db 1777203 CAATGTTAATGACTTGAAGTCAACCTCCCGAGGAAATCTCAACTGCAACACCCCTACT 1777262

QY 463 ACACCTCAATTCAGTAGGAGAGTGTAGAGCAGTGTCTAGCCAACTCCGCCCAACAGTACT 522
Db 1777263 ATGCCCAATTCAGTAGGAGAGTGTAGAGCAGTGTCTAGCCAACTCCGCCCAACAGTACT 1777322

QY 523 TGGGTTTCTCTGTGTAGAGGGTGGACTGAGAGACAGGACTAGCTGATTTCTCTAGGCTGA 582
Db 1777323 TGGGTTTCTCTGTGTAGAGGGGACTGAGAGACAGGACTAGCTGATTTCTCTAGGCTGA 1777382

QY 583 CTAAAGAAATCCNAAAGCTTANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTG 642
Db 1777383 CTAAAGAAATCCNAAAGCTTANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTG 1777442

QY 643 CAATTTAGTCTCACACCCGACCAATCAGAGAGTCTCAATAATGCTTAATCAGGCAAAAACA 702
Db 1777443 CAATTTAGTCTCACAC---CCAAACAGAGAGTTCACCTAAATGCTTAATGAGCAAAAATA 1777498

QY 703 GAGGTTAAAGCAATAGCCCAATCATCTATTGCTGAGAGACAGCAGCGGAGGACAAGGAT 762
Db 1777499 GAGGTTAAAGCAATAGCCCAATCATCTATTGCTGAGAGACAGCAGCGGAGGACAAGGAT 1777558

QY 763 GCGATATAAACTCAGGCAATCAAGCCAGACAAAGCAACCCCTTTGGGTCCCTCCCAATT 822
Db 1777559 AGGATATAAACCCAGGCAATCAGAGCTGGCAACCGCAACCCCTTTGGGTCCCTCCCTTT 1777618

QY 823 GTATGGAGGCTCTGTTTCTACTCTATTCTACTCTATTAAATCATGCAATGCA 875
Db 1777619 GTATGGAGGCTCTGTTTCTCA--CTGTTTCACTCTATTAAATCTTGAACATGCA 1777669

```

```

RESULT 37
US-10-276-774-678
; Sequence 678, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18

```

```

; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 678
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-678

Query Match      35.3%; Score 468.8; DB 17; Length 2052;
Best Local Similarity 88.5%; Pred. No. 1.1e-134;
Matches 533; Conservative 0; Mismatches 59; Indels 10; Gaps 2;

QY 257 CCCTGTATCTTCAACTTCTTGTAAAGTTTGTCTCTTCCAGAAATGCAAGCTCTAAAGCTA 316
Db 993 CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAGCTCTGTAACAA 1052

QY 317 CAAATAGTCTTCAAAATGGAACCCAGATGCAAGTCCATGACTTAAATCTTACCGTGGACCC 376
Db 1053 CACATCGTCTTCAAAATGGAACCCAGATGCAAGTCCATGACTTAAATCTTACCGGATCC 1112

QY 377 CTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACTTGAAGTCAACCCCTCCCGAG 436
Db 1113 CTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACTTGAAGTCAACCCCGAG 1172

QY 437 GAAATCTCACTGCAACCCCTACTACTACCTCAATTCAGTAGGAGGAGTGTAGAGCACT 496
Db 1173 GAAATCTCAATGCAACCCCTACTACTATGCCCCCAATTAAGCAGGAGGAGTGTGGAGCAGT 1232

QY 497 TGCTAGCCCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGAC 556
Db 1233 CGACGGCCCAACCTCCCAACAGCAGTGGGTTTTCTGTTGAGAGGGGTACTGAGAGAC 1292

QY 557 AGGACTAGCTGGATTTCTTCTAGGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGACC 616
Db 1293 AGGACTAGCTGGATTTCTTCTAGGCTGACTAAGAAATCCCTAAGCTTACTGCGGAGGTGACT 1352

QY 617 GCATCATCTTTAAACATGGGGCTTGCACCTTACTGCTCACACCCGACCAATC----- 667
Db 1353 GTATCCACCTTTAAACATGGGGCTTGCACCTTACTGCTCACACCCGACCAATCAGGTAGTAA 1412

QY 668 AGAGAGCTCACTTAAATGCTTAATCAGGCAAAAACAGAGGTAAAGCAATAGCAATCATC 727
Db 1413 AGAGAGCTCACTTAAATGCTTAATAGGCAAAAACAGAGGTAAAGCAATAGTAAAT-ATC 1471

QY 728 TATTGCTGAGAGCAGCGGGAAGGACAAAGGATTGGGATATATAAACTCAGGCAATTCAGGC 787
Db 1472 TATCAGCTGAGAGTACAGGGGAGGAGCAATGATTGGGATAGAAACCCAGGCAATTCGAGC 1531

QY 788 CAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 847
Db 1532 CGGCAACGGCAACCCCTTTGGGTCTCTCTTCCATTTATGGGAGGCTCTGTTTTCCTCACTA 1591

QY 848 TT 849
Db 1592 TT 1593

```

```

RESULT 38
US-10-632-793-28
; Sequence 28, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793

```

Fri Feb 25 16:26:28 2005

```
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-632-793-28

Query Match      35.0%; Score 465; DB 17; Length 3372;
Best Local Similarity 89.7%; Pred. No. 2.2e-133;
Matches 498; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 321 TAGTCTTCAAAATGGAACCCAGATGCGATGCACTAAATCTACGGTGAGCCCTGG 380
Db |||||
Qy 2797 TAAAACTACAATGGAGCCCAAGATGCGATGCAAGACTTAAGATCTACCGCAGACCCCTGG 2856
Db |||||
Qy 381 ACCGCCCTGTAGACTATGCTCTGATGTTAATGACATTAAGTCAACCCCTCCCGAGGAAA 440
Db |||||
Qy 2857 ACCGCCCTGTAGCCCAAGATCTGATGTTAATGACATTAAGTCAAAAGGCAACCCCTCTCGAGGAAA 2916
Db |||||
Qy 441 TCTCAACTGCAACAAACCCCTACTACACTCCAATTCAGTAGGAGCAGTTAGAGCAGTTGTC 500
Db |||||
Qy 2917 TCTCAGCTGCAACACTCTACTAGCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGTC 2976
Db |||||
Qy 501 AGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGAGGGTGAGTGAAGCAGGA 560
Db |||||
Qy 2977 GGCCAACTCCCAACAGCAGTACTAGTTTCTGTTGAGATGGGGAGCTGAGAGCAGGA 3036
Db |||||
Qy 561 CTAGCTGGAATTCCTAGGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGACCGCAT 620
Db |||||
Qy 3037 CTAGCTGGAATTCCTAGGCTGATTAAGAAATCCCTAAGCTAGCTGGGAAGGTGACCAT 3096
Db |||||
Qy 621 CCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTA 680
Db |||||
Qy 3097 CCACCTTTAAACACGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTA 3156
Db |||||
Qy 681 AAATCTTAATCGGCAAAACAGGAGTTAAGCAATAGCAATAGCAATCTATTGCGCTGAGAG 740
Db |||||
Qy 3157 AAATGCTAATTAGGCAAAACAGAGAGGTAAGAAATAGCAATCAATTAATTGCGCTGAGAG 3216
Db |||||
Qy 741 CACAGCGGAAGCAAGGATTTGGATATATAACTCAGGCATTTCAAGCCAGCAACAGCAAC 800
Db |||||
Qy 3217 CACAGCAGGAGGACANTGATCGGATATTAACCAAGTTTTCGAGCCGCAACGGCAAC 3276
Db |||||
Qy 801 CCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTCACTCTATTCTACTATTA 860
Db |||||
Qy 3277 CCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTCACTATTTCATCTATTATTA 3336
Db |||||
Qy 861 AATCATGCAACTGCA 875
Db |||||
Qy 3337 AATCTTGCAACTGCA 3351
Db |||||
```

```
RESULT 39
US-09-997-722-148
; Sequence 148, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 22436
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-997-722-148

Query Match      35.0%; Score 465; DB 11; Length 22436;
Best Local Similarity 88.8%; Pred. No. 6e-133;
Matches 539; Conservative 0; Mismatches 57; Indels 11; Gaps 3;

Qy 257 CCCTGTATCTTCAACTTCCTTGAAGTTTGTCTCTCCAGAAATCGAAGCTGTAAAGCTA 316
Db |||||
Qy 5655 CCCTGTACCTTTAACTCTCTTGAAGTTTGTCTCTCCAGAAATCGAAGCTGTAAAGCTA 5714
Db |||||
Qy 317 CAAATAGTTCTTCAAAATGGAACCCAGATGCGATGCCATGATCAAAATCTACCGTGGACCC 376
Db |||||
Qy 5715 CAAATGGTCTTCAAAATGGAACCCAGATGCGATGCCATGATCAAGATCTTCTGAGACCC 5774
Db |||||
Qy 377 CTGAGCGGCTGCTAGACT-ATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGA 435
Db |||||
Qy 5775 CTGAGCTGGCTGCTAGCTCTATGCTCCAAATGTTAATGACATTTGAAGGCAACTCTCCCGA 5834
Db |||||
Qy 436 GGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGGAGCAGTTAGAGCAG 495
Db |||||
Qy 5835 GGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGGAGCAGTTAGAGCAA 5894
Db |||||
Qy 496 TTGTGAGCCAACTCCCCCAACAGTACTTTGGGTTTCTGTTGAGAGGGTGGACTGAGAGA 555
Db |||||
Qy 5895 TCATTGGCCAACTCCCCCAACAGCAGCTTTGGGTTTCTGTTGAGAGGGGAGCTGAGAG- 5953
Db |||||
Qy 556 CAGGACTAGCTGATTTCTAGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGAC 615
Db |||||
Qy 5954 CAGGACTAGCTGATTTCTAGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGAC 6013
Db |||||
Qy 616 CGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 667
Db |||||
Qy 6014 TGACCCACCTTTAAACACGGGCTTGCAACTTAGTTTACACCTGACCAATCAAGTAGTA 6073
Db |||||
Qy 668 -AGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGAGAGTTAAGCAATAGCCAAATCAT 726
Db |||||
Qy 6074 AAGAGAGCTCACTAAATGCTAATTAGGCAAAAACAGAGAGTTAAGCAATAGCCAAATCAT 6133
Db |||||
Qy 727 CTATTGCTGAGACACAGCGGAGCAAGGATTTGGGATATATAAACTCAGGCAATTCAG 786
Db |||||
Qy 6134 CTATCCTGAGAGCAGCGGTGGGAGCAATGATCAGGATATTAACCCAGGCAATTTGAG 6193
Db |||||
Qy 787 CCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGATGGGAGCTCTGTTTCACTCT 846
Db |||||
Qy 6194 CCAGCAATGCTACCTCTTTGGTCCCTCCCTTTGATGGGAGCTCTGTTTCACTCT 6253
Db |||||
Qy 847 ATTTCACT 853
Db |||||
Qy 6254 ATTAAC 6260
Db |||||
```

```
RESULT 40
US-10-087-192-1666
; Sequence 1666, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1666
; LENGTH: 285020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(285020)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1666

Query Match      34.6%; Score 460; DB 13; Length 285020;
Best Local Similarity 88.6%; Pred. No. 8.5e-131;
Matches 535; Conservative 0; Mismatches 57; Indels 12; Gaps 3;

Qy 257 CCCTGTATCTTCAACTTCCTTGAAGTTTGTCTTCCAGAAATTGAAGCTGTAAAGCTA 316
Db 279486 CCCTGTATCTTCAACTTCCTTGAAGTTTGTCTTCCAGAAATTGAAGCTGTAAAGCTA 279545

Qy 317 CAAATAGTTCTTCAATGAACCCAGATGCAGTCCATGACTTAAATCTACCGTGACCC 376
Db 279546 CAAATAGTTCTTCAATGAACCCAGATGCAGTCCATGACTTAAATCTACCGTGACCC 279605

Qy 377 CTGGACCGCCTGTAGACTATGCTCTGATGTATGATTAATGACATTGAAGTCAACCCCTCCCGAG 436
Db 279606 CTGGACCGCCTGTAGACTATGCTCTGATGTATGATTAATGACATTGAAGTCAACCCCTCCCGAG 279665

Qy 437 GAAATCTCAACTGCACCAACCCCTACTACACTCAATTCAGTAGGAGCAGTTAGAGCAGT 496
Db 279666 GAAATCTCAACTGCACCAACCCCTACTACACTCAATTCAGTAGGAGCAGTTAGAGCAGT 279725

Qy 497 TGTACGCCAACCTCCCAACAGTACTTGGGTTTTCTGTGTGAGAGGTTGGACTGAGAGAC 556
Db 279726 CATTGSCCAACCTCCCAACAGCAGTCTGGGTTTTCTGTGTGAGAGTGGGAGCTGAGAGAC 279785

Qy 557 AGGACTAGCTGATTTCTTAGCTAGCTGACTAAG--AATCCCNAAAGCTTANCTGGGAGGTGA 614
Db 279786 AGGACTAGCTGATTTCTTAGCTAGCTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGA 279845

Qy 615 CCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGCCCAATC----- 667
Db 279846 CCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGCCCAATCAGGAGT 279905

Qy 668 --AGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGAGGTAAAGCAATAGCCCAATCA 725
Db 279906 AAAGACAGCTCACTAAATGCTAATTAGGCTAAACAGAGGTAAAGAAATAGCCCAATCA 279965

Qy 726 TCTATTGCTGAGACACAGCGGGAAGGACAGAGATTGGATATTAATCTCAGGCATTCAA 785
Db 279966 TCTATCGCTGAGAGCAC--GAGGGAGGGAATGATCAGGATATAAACCAGGCATTCAA 280024

Qy 786 GCCAGCAACAGAACCCCTTTGGGTCCCTCCATTGTATGGAGCTCTGTTTTCACTC 845
Db 280025 GCCGGCAGTGGCTACCCGCTTTGGGTCCCTCTGTTTGTATGGAGCTCTGTTTTCACTC 280084

Qy 846 TATT 849
Db 280085 TATT 280088
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:52:54 ; Search time 4132.82 Seconds
(without alignments)
12240.400 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatgaagagcttttag.....gcaaggatcccccaataca 1329

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_g881.*
- 9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	706.8	53.2	2748	3	CR605851 full-leng
2	702.8	52.9	2749	3	CR617248 full-leng
3	702	52.8	2500	3	CR622175 full-leng
4	699	52.6	2748	3	CR625046 full-leng
5	684.2	51.5	1071	5	CR635066 BX365066
6	678.6	51.1	2716	5	CR613169 full-leng
7	669.4	50.4	998	5	CR337769 BX337769
8	661.2	49.8	1058	5	CR378303 BX378303
9	652.2	49.1	1019	5	CR439636 BX439636
10	610	45.9	966	5	CR380176 BX380176
11	608.2	45.8	758	5	CR357208 BX357208
12	593.4	44.7	771	4	CR1087886 BX1087886
13	592	44.5	4204	3	CR035153 BX035153
14	561	42.2	719	7	CR272394 BX272394
15	544	40.9	900	5	CR459153 BX459153
16	500.6	37.7	1002	5	CR953822 CRF2781M
17	480.8	36.2	679	9	CR113694 BX113694
18	480.8	36.2	999	5	CR918330 BX918330
19	475.2	35.8	683	9	CR134524 BX134524
20	474	35.7	669	9	CR049953 BX049953
21	473	35.6	681	9	CR116971 BX116971
22	469.6	35.3	710	9	CR096341 BX096341
23	468.2	35.2	1160	3	CR030968 BX030968
24	465.2	35.0	893	5	CR437925 BX437925

25	456.2	34.3	670	9	AG072852 Pan trogl
26	455	34.2	712	9	AG099717 Pan trogl
27	439.2	33.0	689	9	AG121669 Pan trogl
28	438.8	33.0	921	8	BH149565 ENTQ48TR
29	435.6	32.8	679	9	AG076758 Pan trogl
30	433.6	32.6	683	9	AG102951 Pan trogl
31	429.4	32.4	611	8	AQ381711 RPI111-16
32	429.4	32.3	736	9	AG086046 Pan trogl
33	427.6	32.2	714	8	AQ238712 RPI111-64
34	424.4	31.9	697	9	AG108186 Pan trogl
35	423.8	31.9	675	9	AG117095 Pan trogl
36	422.2	31.8	797	6	CB990962 AGENCOURT
37	422.2	31.8	803	6	CB992422 AGENCOURT
38	422.2	31.8	826	6	CB990802 AGENCOURT
39	422.2	31.8	837	6	CB996320 AGENCOURT
40	422.2	31.8	852	6	CB994833 AGENCOURT
41	419.4	31.6	672	9	AG040241 Pan trogl
42	418.8	31.5	611	7	CR542951 DKF29470A
43	418.2	31.5	870	7	CR645411 LILUMIGEN
44	416.2	31.3	692	9	AG114407 Pan trogl
45	415.6	31.3	494	1	AA781423 aj26c03.s

ALIGNMENTS

RESULT 1
CR605851
LOCUS full-length cDNA clone CS0DE012YJ24 of Placenta of Homo sapiens (human).
DEFINITION
ACCESSION CR605851
VERSION CR605851.1 GI:50486658
KEYWORDS HTC; CDS; CDS; CDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2748)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 2748)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
source Location/Qualifiers
..2748 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="placenta"
/plasmid="pCMVSPORT_6"

Query Match 53.2%; Score 706.8; DB 3; Length 2748;
Best Local Similarity 90.5%; Pred. No. 1.2e-202;
Matches 769; Conservative 0; Mismatches 69; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTACCGCCAAAGAGGGGACCTGTTATTTT 60
|||||
DB 1911 TCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGGGGACCTGTTATTTT 1970
|||||

61 AGGGGAAGAAGCTGTTAGTATGTTAAATCAATCTGGAATCAATTAAGTGAAGAAGTTAAAGA 120
1971 AGGGGAAGAAGCTGTTATGTTAAATCAATCGGAATCGTCACTGAGAAGAAGTTAAAGA 2030
121 AATTGAGATCGAATATATATAGTAGAGAGAGAGCTTCAAAACACAGTGCACCTGGGCGCT 180
2031 AATTGAGATCGAATATCAACGCTAGAGAGAGAGCTTCCGAAACACAGTGCACCTGGGCGCT 2090
181 CCTCAGCCAAATGAGTGCCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
2091 CCTCAGCCAAATGAGTGCCTGAGTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 2150
241 TTACTCTCTTTGGACCCCTGATCTTCAACTCTCTGTTAAGTTGCTCTTCCAGAAAT 300
2151 GCTACTCTCTTTGGACCCCTGATCTTAAACCTCTCTGTTAAGTTGCTCTTCCAGAAAT 2210
301 TGAAGCTGTAAGCTTACAATAGTCTTCAATGGAACCCAGATGCGAGTCCATGACTAA 360
2211 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCGAGTCCATGACTAA 2258
361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 420
2259 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCA 2318
421 AGTCACCCCTCCCGAGGAATCTCACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
2319 AGGCACCCCTCCCTGAGGAATCTCAGCTGCAACCTCTACTACGCCCCCAATTCAGCAG 2378
481 AAGCAGTTAGAGAGTGTGACGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
2379 AAGCAGTTAGAGAGTGTGCGGCAACCTCCCAACAGCATTAGGTTTCTCTGTTGAGA 2438
541 GGGTGGACTGAGAGACAGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCNAAAGCT 600
2439 TGGGGACTGAGAGACAGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCNAAAGCT 2498
601 ANCTGGGAAGGTGACCGCATTCATTTTAAACATGGGGCTTCCAACTAGCTCAGACCG 660
2499 AGCTGGGAAGGTGACCAATCCACCTTTTAAACACGGGGCTTGCAACTTAGCTCAGACCTG 2558
661 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
2559 ACCAATCAGAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCC 2618
721 AATCATCTATTGCTGAGAGCAGCGGAAGCAAGGATTTGGGATATAAATCAGCA 780
2619 AATCATCTATTGCTGAGAGCAGCGAGGAGCAATGATCGGGATATAAATCCCAAGTC 2678
781 TTCAGCCAGCAACGACCAACCCCTTTGGTCCCTCCCATTTGTATGGGAGCTCTGTTTT 840
2679 TTCAGCCGCAACGCAACCCCTTTGGTCCCTCCCTTTGTTGATGGGAGCTCTGTTTT 2738
841 CACTCTATT 850
2739 CATGCTATT 2748

RESULT 2
LOCUS CR617248
DEFINITION full-length cDNA clone CS0D1022YJ18 of Placenta Cot 25-normalized of Homo sapiens (human).
ACCESSION CR617248
VERSION CR617248.1 GI:50498055
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2749)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL
REMARK
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2749)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. 2749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1022YJ18"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 52.9%; Score 702.8; DB 3; Length 2749;
Best Local Similarity 90.4%; Pred. No. 1.9e-201;
Matches 765; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
ORIGIN
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTATTTT 60
DB 1916 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 1975
QY 61 AGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCATTAATCAGAAAGTTAAAGA 120
DB 1976 AGGGAAGAATGCTGTTATGTTAATCAATCTGGAATCGTCACTGAGAAAGTTAAAGA 2035
QY 121 AATTGAGATCGAATATATATAGTAGAGAGAGCTTCAAAACACACTGCACCTCGGCGCT 180
DB 2036 AATTGAGATCGAATATCAACGCTAGAGAGAGAGCTTCAAAACACTGCACCTCGGCGCT 2095
QY 181 CCTCAGCAATGAGATGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
DB 2096 CCTCAGCAATGAGATGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 2155
QY 241 TTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTTAAAGTTTGTCTTCCAGAAAT 300
DB 2156 GCTACTCTCTTTGGACCCCTGATCTTCAACTCTCTTAAAGTTTGTCTTCCAGAAAT 2215
QY 301 TGAAGCTGTAAGAGTACAAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 360
DB 2216 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCGAGTCCATGACTAA 2263
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTGA 420
DB 2264 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACTCAA 2323
QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
DB 2324 AGGCACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 2383
QY 481 AAGCAGTTAGAGAGTGTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
DB 2384 AAGCAGTTAGAGAGTGTGTCGCGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 2443
QY 541 GGGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCT 600
DB 2444 TGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCT 2503
QY 601 ANCTGGGAAGGTGACCGCATTCATTTTAAACATGGGGCTTGCACACTTAGCTCACACCG 660
DB 2504 AGCTGGGAAGGTGACCAATCCACCTTTTAAACAGGGGCTTGCACACTTAGCTCACACCTG 2563

```

QY 661 ACCAATCAGAGGCTCAATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2564 ACCAATCAGAGGCTCAATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2623

QY 721 AATCATCTATTGCTTGGAGGACAGCAGCGGAGGACAGGATTTGGGATATAAATCAGGCA 780
Db 2624 AATCATCTATTGCTTGGAGGACAGCAGCGGAGGACAGGATTTGGGATATAAATCAGGCA 2683

QY 781 TCAAGCAGGACAGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTT 840
Db 2684 TTCGAGCGGCAACGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTT 2743

QY 841 CACTCT 846
Db 2744 CATGCT 2749

RESULT 3
LOCUS CR622175 2500 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR622175
VERSION CR622175.1 GI:50502982
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2500)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
REMARK - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1..2500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 52.8%; Score 702; DB 3; Length 2500;
Best Local Similarity 90.6%; Pred. No. 3.2e-201;
Matches 763; Conservative 0; Mismatches 67; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTAACTGGAATCAATCTGGAATCTTAAAGAGTTAAAGA 120
Db 1669 TCAAAATCGAAGAGCTTTAGACTTGTAACTGGAATCAATCTGGAATCTTAAAGAGTTAAAGA 1728

QY 61 AGGGAGAGATGCTGTTAGTATGTTAATCAATCTGGAATCTTAAAGAGTTAAAGA 120
Db 1729 AGGGAGAGATGCTGTTAGTATGTTAATCAATCTGGAATCTTAAAGAGTTAAAGA 1788

QY 121 AATTTCAGATCGAATATAATGTAGAGCAGAGGACCTTCAAACAACCTGCACCTGGGCGCT 180
Db 1789 AATTTCAGATCGAATATAACAGTAGAGCAGAGGAGCTTCGAAACACTGGACCTGGGCGCT 1848

```

```

QY 181 CCTCAGCAATGGATGCGCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 1849 CCTCAGCAATGGATGCGCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1908

QY 241 TTTTACTCTTTTGGAGCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTCCAGAAAT 300
Db 1909 GCTACTCTCTTTTGGAGCCCTGTATCTTAACTCTCTTGTAACTTTGTCTCTCCAGAAAT 1968

QY 301 TGAAGCTGTAAAGCTCAAAATAGTTCTTCAAAATGGAAACCCAGATGCGAGTCAATGACTAA 360
Db 1969 CGAAGCTGTAAACTA-----CAAATGGAGCCCAAGATGCGAGTCAAGACTAA 2016

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTGATGTTAATGACATTGA 420
Db 2017 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCA 2076

QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGG 480
Db 2077 AGGCACCCCTCTGAGGAAATCTGCTGTGCAACCTCTACTACGCCCAATTCAGCAGG 2136

QY 481 AAGCAGTTTAGAGCAGTTGCTCAGCCAACTCCCAACAGTACTTTGGGTTTTCTGTGAGA 540
Db 2137 AAGCAGTTTAGAGCGTCTGCGCCAACTCCCAACAGTACTTTGGGTTTTCTGTGAGA 2196

QY 541 GGGTGGACTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTGAATCCCAAGCCT 600
Db 2197 TGGGGGACTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTGAATCCCAAGCCT 2256

QY 601 ANCTGGGAGGTGACCGCATCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCG 660
Db 2257 AGCTGGGAGGTGACCATCCCTTTAAACATGGGCTTGCAACTTAGCTCACACCTG 2316

QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2317 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2376

QY 721 AATCATCTATTGCTTGGAGCAGCAGCGGAGGACAGGATTTGGATATAAATCAAGTC 780
Db 2377 AATCATCTATTGCTTGGAGCAGCAGCGGAGGACAGGATTTGGATATAAATCAAGTC 2436

QY 781 TTCAGGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGGCTCTCTTTT 840
Db 2437 TTCGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGGCTCTCTTTT 2496

QY 841 CA 842
Db 2497 CA 2498

RESULT 4
LOCUS CR625046 2748 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI044YK06 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR625046
VERSION CR625046.1 GI:50505853
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2748)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

```

COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source
1..2748
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1044YK06"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 52.6%; Score 699; DB 3; Length 2748;
Best Local Similarity 90.6%; Pred. No. 2.7e-200;
Matches 760; Conservative 0; Mismatches 67; Indels 12; Gaps 1;
Qy 1 TCAAAATCGAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACTGTTATTTT 60
Db 1922 TCAAAATCGAGAGCTTTAGACTTCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 1981
Qy 61 AGGGGAAGATGCTGTTAGTATGTTAAATCAATCTGGAATCAATCTAGAGAAAGTTAAAGA 120
Db 1982 AGGGGAAGATGCTGTTATTTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTAAAGA 2041
Qy 121 AATTGAGATCGAATATAATGATGAGCAGAGGACCTTCAAAACACTGCACCTCGGGGCT 180
Db 2042 AATTCGAGATCGAATACACGCTAGAGCAGAGGAGCTTCGAAACACTGCACCTCGGGGCT 2101
Qy 181 CCTAGCAATGAGATGCCCTGAGACTCTCCCTCTTAGAGCTCTAGCAGCTATAATTT 240
Db 2102 CCTAGCAATGAGATGCCCTGAGACTCTCCCTCTTAGAGCTCTAGCAGCTATAATTT 2161
Qy 241 TTACTCTCTTTGGACCTGATCTTCAACTCTCTTGAAGTTGCTCTTCAGAAAT 300
Db 2162 GCTACTCTCTTTGGACCTGATCTTAACTCTCTGTTAACTTTGCTCTTCAGAAAT 2221
Qy 301 TGAAGCTTAAGCTTACAAATAGTTCTTTCAAATGGAACCCAGATGCACTGACTGACTAA 360
Db 2222 CGAAGCTGTAATACTA-----CAATGGAGCCCAAGATGCACTGCAAGACTAA 2269
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTGA 420
Db 2270 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACTCAA 2329
Qy 421 AGTCAACCCCTCCGAGGAAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGG 480
Db 2330 AGGCACCCCTCTGAGGAATCTAGCTGCAACACCTCTACTACGCCCCCAATTCAGCAGG 2389
Qy 481 AAGCAGTTAGACGAGTTGTACGCCAACCTCCCAACAGTACTTTGGGTTTTCTGTTGAGA 540
Db 2390 AAGCAGTTAGACGAGTGTGTCGCCAACCTCCCAACAGCACTTAGTTGTTTCTGTTGAGA 2449
Qy 541 GGGTGGACTGAGACAGCAGTACTGATTTCTTAGGCTGACTAAGATCCCAAGCCT 600
Db 2450 TGGGGGACTGAGACAGCAGTACTGATTTCTTAGGCTGACTAAGATCCCTAAGCCT 2509
Qy 601 ANCTGGGAAGGTGACCGCATCATCTTTAAACATGGGGCTTGGCACTTAGCTCACACCGG 660
Db 2510 AGCTGGGAAGGTGACCATCATCCACTTTAAACAGGGGCTTGGCACTTAGCTCACACCTG 2569
Qy 661 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAAAAGGAGGTAAGCAATAGCC 720
Db 2570 ACCAATCAGAGCTCACTAAATGCTAATTAGGCAAGACAGGAGGTAAGCAATAGCC 2629
Qy 721 AATCATCTATTGCTGAGACACGCGGAGGACAGGATGGGATATAACTCAGCA 780
Db 2630 AATCATCTATTGCTGAGACACGAGGAGGACAAATGATCGGGATATAAACCAAGTC 2689
Qy 781 TTCAAGCCGCAACAGCAACCCCTTTGGGTCCCCCTCCCATTTGTTATGGGAGCTCTGTTT 839

Db 2690 TTCAGGCGGACAGGGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTT 2748
RESULT 5
BX365066/c
LOCUS
DEFINITION BX365066 Homo sapiens PLACENTA Cot 25-NORMALIZED Homo sapiens cDNA
clone CS0D1046YA18 3-PRIME, mRNA sequence.
ACCESSION BX365066
VERSION BX365066.2 GI:46304105
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1071)
AUTHORS Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30374869.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS1A1012ZC10NP1&c=4215.r.
FEATURES
source
1..1071
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1046YA18"
/tissue_type="PLACENTA Cot 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA Cot 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 51.5%; Score 684.2; DB 5; Length 1071;
Best Local Similarity 89.9%; Pred. No. 6.4e-196;
Matches 755; Conservative 3; Mismatches 69; Indels 13; Gaps 2;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCCAAAGAGGGGAACTGTTATTTT 60
Db 827 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAAAGAGGGGAACTGTTATTTT 768
Qy 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCAATCTAGAAAGTTAAAGA 120
Db 767 AGGGGAAGATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 708
Qy 121 AATTGAGATCGAATATAATGATGAGCAGAGGACCTTCAAAACACTGCACCTCGGGGCT 180
Db 707 AATTCGAGATCGAATACACGCTAGAGCAGAGGAGCTTCGAAACACTGCACCTCGGGGCT 648
Qy 181 CCTCAGCAATGAGATGCCCTGAGACTCTCCCTCTTAGGACCTTAGCAGCTATAATTT 240
Db 647 CCTCAGCAATGAGATGCCCTGAGACTCTCCCTCTTAGGACCTTAGCAGCTATAATTT 588
Qy 241 TTACTCTCTTTGGACCTGATCTTCAACTCTCTTGAAGTTGCTCTTCAGAAAT 300
Db 587 GCTACTCTCTTTGGACCTGATCTTTAACTCTCTTGAAGTTGCTCTTCAGAAAT 528
Qy 301 TGAAGCTGTTAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCACTGACTGACTAA 360

Db 527 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCAGTCCCAAGACTAA 480

Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420

Db 479 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 420

Qy 421 AGTCACCCCTCCCGAGGAATCTCAACTGACAAACCCCTACTACACTCCAAATTCAGTAGG 480

Db 419 AGGCACCCCTCCCGAGGAATCTCAGCTGCACAACTCTTACTACGCCCCCAATTCAGCAGG 360

Qy 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTAGTCTTGGGTTTTTCCCTGTTGAGA 540

Db 359 AAGCAGTTAGAGCGGTGCTGGCCCACTCCCAACAGCAGTCTAGGTTTTTCCCTGTTGAGA 300

Qy 541 GGGTGAGCTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCT 600

Db 299 TGGGGGACTGAGAGACAGACTAGCTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCT 240

Qy 601 ANCTGGGAAGTGCACCGCATCTTTAAACATGGGCTTGCACCTTAGCTCACACCCG 660

Db 239 AGCTGGGAAGTGCACCATCTCCACCTTTAAACAGGGGCTTGCACCTTAGCTCACACCTG 180

Qy 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 720

Db 179 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 120

Qy 721 AATCATCTATTGCTGAGACACAGCGGAGACAAAGATTTGGGATATATAAATCAGGCA 780

Db 119 AATCATCTATTGCTGAGACACAGCGGAGGACAAATGATCGGGATATATAAATCAGGCA 60

Qy 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGCTCCCTCCCACTTATGAGGAGCTCTGTTTT 840

Db 59 TTCAGCGCGCAACGGCAACCCCTTTGGTNNCCCTCCC-TTGTATGCGASMTCKGTTTT 1

RESULT 6
LOCUS CR613169 2716 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DE013Y120 of Placenta of Homo sapiens (human).

ACCESSION CR613169
VERSION CR613169.1 GI:50493976
KEYWORDS HTC; CNSLT_cDNA
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
2 (bases 1 to 2716)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
Location/Qualifiers
1. 2716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013Y120"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

RESULT 7
BX337769/c 998 bp mRNA linear EST 07-APR-2004
LOCUS BX337769 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI051YM13 3-PRIME, mRNA sequence.
ACCESSION BX337769
VERSION BX337769.2 GI:46272079
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

ORIGIN
Query Match 51.1%; Score 678.6; DB 3; Length 2716;
Best Local Similarity 90.5%; Pred. No. 4.3e-194;
Matches 739; Conservative 0; Mismatches 66; Indels 12; Gaps 1;
Qy 1 TCAAAATCGAAGAGCTTTAGAGCTTCTCTAAACCCCAAAAGAGGGGAACTCTGTTTATTTT 60
Db 1912 TCAAAATCGAAGAGCTTTAGAGCTTCTCTAAACCCCTGAAGAGGGGAACTCTGTTTATTTT 1971
Qy 61 AGGGGAAGAATCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
Db 1972 AGGGGAAGAATCTGTTATTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 2031
Qy 121 AATTTGAGATCGAATATATATGATAGACAGAGGACCTTCAAAACACTGCAACCCCTGGGCGCT 180
Db 2032 AATTCGAGATCGAATACAACTGATAGACAGAGAGCTTCGAAACACTGGACCTCGGCGCT 2091
Qy 181 CCTCAGCCAATGGAATGCTGCTCTCCCTTTCTTTAGGACCTCTAGCAGCTATATATTT 240
Db 2092 CCTCAGCCAATGGAATGCTGCTCTCCCTTTCTTAGGACCTCTAGCAGCTATATATTT 2151
Qy 241 TTTTACTCTCTTTGGACCTCTGATCTTCAACTCTCTTGTAGTATGTTTGTCTCTTCCAGAAT 300
Db 2152 GCTACTCTCTCTTTGGACCTCTGATCTTTAACTCTCTTGTAACTTTGTCTCTTCCAGAAT 2211
Qy 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db 2212 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCAGTCCCAAGACTAA 2259
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2260 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 2319
Qy 421 AGTCACCCCTCCCGAGGAATCTCAACTGACAAACCCCTACTACACTCCAATTCAGTAGG 480
Db 2320 AGGCACCCCTCCCGAGGAATCTCAGCTGCACAAACCTCTACTACGCCCCCAATTCAGCAGG 2379
Qy 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTTTCTGTTGAGA 540
Db 2380 AAGCAGTTAGCGGTCGTCGGCCCACTCCCAACAGACACTTAGGTTTTTCTGTTGAGA 2439
Qy 541 GGGTGGAGCTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCT 600
Db 2440 TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCT 2499
Qy 601 ANCTGGGAAGTGCACCGCATCTTAAACATGGGCTTGCACCTTAGCTCACACCCG 660
Db 2500 AGCTGGGAAGTGCACCATCTCCACCTTTAAACACGGGGCTTGCACCTTAGCTCACACCTG 2559
Qy 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 720
Db 2560 ACCAATCAGAGAGCTCACTAAATGCTTAATGAGCAAGACAGAGGTAAAGCAATAGCC 2619
Qy 721 AATCATCTATTGCTGAGAGCACAGCGGAGGACAAAGGATTCGGATATATAAATCAGGCA 780
Db 2620 AATCATCTATTGCTGAGAGCACAGCGGAGGACAAATGATCGGGATATATAAATCAGGCA 2679
Qy 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGCTCCCTCCC 817
Db 2680 TTCAGCGCGCAACGGCAACCCCTTTGGGCTCCCTCCC 2716

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 998)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30337641.

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1051AG07NP1&c=4215.r.

FEATURES
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized." source

ORIGIN
Query Match 50.4%; Score 669.4; DB 5; Length 998;
Best Local Similarity 89.2%; Pred. No. 2e-191;
Matches 754; Conservative 2; Mismatches 74; Indels 15; Gaps 3;

QY 1 TCAAAATCGAAGCTTTAGACTTCTGCTAATCGGCAAGAGGGGAACTGTTATTTT 60
DB 835 TCAAAATCGAAGCTTTAGACTTCTGCTAATCGGCAAGAGGGGAACTGTTATTTT 776
QY 61 AGGGAAGAACTCTGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAGTTAAAGA 120
DB 775 AGGGAAGAACTCTGTATGTTAATCAATCGGCAAGAGGGGAACTGTTATTTT 716
QY 121 AATTTGAGATCGAATATATATAGACGAGAGAGACCTTCAAAACATGCACTCGGGGCT 180
DB 715 AATTCGAGATCGAATACACGATAGACGAGAGAGCTTCGAAACACTGGACCTGGGGCT 656
QY 181 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 655 CCTCAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
QY 241 TTTACTCTCTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 595 GCTACTCTCTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 536
QY 301 TGAAGCTGTAAGCTACAAATAGTCTTCAATGGAACCCCAAGATGCACTCAATGACTAA 360
DB 535 CGAAGCTGTAAGCTACAAATAGTCTTCAATGGAACCCCAAGATGCACTCAATGACTAA 488
QY 361 AATCTACCTGGACCCCTGGACCCCTGGACCCCTGGACCCCTGGACCCCTGGACCCCTGGAC 420
DB 487 GATCTACCTGGACCCCTGGACCCCTGGACCCCTGGACCCCTGGACCCCTGGACCCCTGGAC 428
QY 421 A-GTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTCTACTACACTCAATTCAGTAG 479
DB 427 ACGGCAACCCCTCTGAGGAAATCTCAAGTGCACAACTCTACTACGCCCCCAATTCAGCAG 368
QY 480 GAAGCAAGTTAGACGCTTGTACGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAG 539
DB 367 GAAGCAAGTTAGACGCTTGTACGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAG 308

QY 540 AGGTTGGAAGCTGAGACAGAGCTAGCTGGATTTCTTAGCTACTAAGATCCCAAGCC 599
DB 307 ATGGGGAGCTGAGACAGAGCTAGCTGGATTTCTTAGCTACTAAGATCCCAAGCC 248
QY 600 TANTCGGAAGGTGACCGCATCCATTTTAAACATGGGGTTTCCAACTTAGCTCACACCC 659
DB 247 TAGCTGGGAAGGTGACCATCCATCTTTAAACACGGGGCTTGCACACTTGGCTCACACCT 188
QY 660 GACCAATCAGAGAGCTCACTAAATCTATCATGCGCAAAACAGGAGGTAAAGCAATAGC 719
DB 187 GACCAATCAGAGAGCTCACTAAATCTATTAGGCAAAACAGGAGGTAAAGCAATAGC 128
QY 720 CAATCATCTATTCCCTGAGAGCAGCGGAGGACAAAGGATTGGGATATAAACTCAGGC 779
DB 127 CAATCATCTATTCCCTGAGAGCAGCGGAGGACAAATGATCGGGATATAAAACCAAGT 68
QY 780 ATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTT--GTATGGGAGCTCTGT 837
DB 67 CTTGAGCGCGCAACGCGCAACCCCTTTGGGTCCCNCCCTTTGTANTGGAGGTCTGT 8
QY 838 TTTC A 842
DB 7 TTTC A 3

RESULT 8
BX378303/c
LOCUS BX378303 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1018Y07 3-PRIME, mRNA sequence.
BX378303
ACCESSION BX378303.2 GI:46557492
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1058)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30439129.

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1018CD04NP1&c=4215.r.

FEATURES
Location/Qualifiers
1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1018Y07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was
normalized." source

ORIGIN
Query Match 49.8%; Score 661.2; DB 5; Length 1058;
Best Local Similarity 86.5%; Pred. No. 6.2e-189;
Matches 727; Conservative 9; Mismatches 92; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGCTTTAGACTTCTGCTAATCGGCAAGAGGGGAACTGTTATTTT 60

```
Db 832 TCNAATCGAGAGCTTTAGACTTGTACCGCTGAAGAGGKKAWCTTTTTTTTT 773
Qy 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCAATCTAGTGAAGAGTTAAAGA 120
Db 772 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCKKAWTKTCACTGAGAAAAGTTAAAGA 713
Qy 121 AATTTGAGATCGAATATATGATGAGACAGAGGACCTTCAAAACACATGCAACCTGGGGCCT 180
Db 712 AATTCGAGATCGAATATCAAGTATGAGACAGAGGAGCTTCGAAACATCTGGACCTGGGGCCT 653
Qy 181 CTTGAGCAATGATGCTGCTGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 652 CCTGAGCAATGATGCTGCTGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 593
Qy 241 TTTTACTCTCTTTGGACCTGTATCTTCAATCTCTTCTTGAATTTGTCTCTTCCAGAAAT 300
Db 592 GCTACTCTCTTTGGACCTGTATCTTCAATCTCTTCTTGAATTTGTCTCTTCCAGAAAT 533
Qy 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATCGAGTCCATGACTAA 360
Db 532 CGAAGCTGTAAAACTA-----CAATGGAGCCCAAGATCGAGTCCAGACTAA 485
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
Db 484 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 425
Qy 421 AGTCACCCCTCCGAGAGAAATCTCAATGTCACAAACCCCTACTACTACTCCAGTGG 480
Db 424 AGGCACCCCTCTGAGGAAATCTCAGCTGCACAAACCTCTACTACGCCCAATTCAGCAGG 365
Qy 481 AAGCAGTTAGACAGTGTGTCAGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
Db 364 AAGCAGTTAGACAGGCTGCTGGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA 305
Qy 541 GGGTGGACTCAGAGACAGGACTAGCTGGAATTTCTAGGCTGACTAAGAAATCCCNAAAGCCT 600
Db 304 TGGGGNACTCAGAGACAGGACTAGCTGGAATTTCTAGGCTGACTAAGAAATCCCNAAAGCCT 245
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGGCACTTAGCTCACACCG 660
Db 244 AGCTGGGAAGGTGACCATCCACCTTTAAACACGGGGCTTGGCACTTAGCTCACACCTG 185
Qy 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGCAGCAAAACAGAGGTAAGCAATAGCC 720
Db 184 ACCAATCAGAGAGCTCACTAAATGCTAATCAGCAGCAAAACAGAGGTAAGCAATAGCC 125
Qy 721 AATCATCTATTGCTGAGAGACAGCGGGAAGGACAGGATGGGATATATAAATCAGGCA 780
Db 124 AATCATCTATTGCTGAGAGACAGCGGGAAGGACAGGATGGGATATATAAATCAGGCA 65
Qy 781 TTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGGAGCTCTGTTTT 840
Db 64 TTCAGCGCGCAACGGCAACCCNNTNTGTTNCCNTGTTGTTGTTGTTGTTGTTT 5
```

```
RESULT 9
BX439636/c 1019 bp mRNA linear EST 04-MAY-2004
LOCUS BX439636 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012Y24
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX439636
VERSION BX439636.2 GI:47000005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1019)
Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30771765.
```

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: secrefegenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?cs=CS0DE012BE12NP16c=4215.r>.
Location/Qualifiers
1..1019
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012Y24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

FEATURES

source

ORIGIN

```
Query Match 49.1%; Score 652.2; DB 5; Length 1019;  
Best Local Similarity 85.5%; Pred. No. 3.3e-186;  
Matches 725; Conservative 19; Mismatches 91; Indels 13; Gaps 2;  
Qy 1 TCNAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAAAGAGGGGAACTGTTATTTTT 60  
Db 835 TCNAATCGAAGAGCTTTAGACTTGTCTAACCCGCTGAAGAGGGGAACTGTTATTTTT 776  
Qy 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCAATCTAGGAAAGTTAAAGA 120  
Db 775 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 716  
Qy 121 AATTTGAGATCGAATATATGATGAGACAGAGGACCTTCAAAACACATGCAACCTGGGGCCT 180  
Db 715 AATTCGAGATCGAATATCAAGTATGAGACAGAGGAGCTTCGAAACATCTGGACCTGGGGCCT 656  
Qy 181 CTTGAGCAATGATGCTGCTGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240  
Db 655 CCTGAGCAATGATGCTGCTGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 596  
Qy 241 TTTTACTCTCTTTGGACCTGTATCTTCAATCTCTTGTAGTATGTTGTTGTTTCTTCCAGAAAT 300  
Db 595 GCTACTCTCTTTGGACCTGTATCTTCAATCTCTTGTAGTATGTTGTTGTTTCTTCCAGAAAT 536  
Qy 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATCGAGTCCATGACTAA 360  
Db 535 CGAAGCTGTAAAACTA-----CAATGGAGCCCAAGATCGAGTCCAGACTAA 488  
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420  
Db 487 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 428  
Qy 421 AGTCACCCCTCCGAGAGAAATCTCAATGTCACAAACCCCTACTACTACTCCAAATTCAGTAGG 480  
Db 427 AGGCACCCCTCTCTGAGGAAATCTCAGCTGCMACAACTCTATACCGCCCAATTCAGCAGG 368  
Qy 481 AAGCAGTTAGACAGTGTGTCAGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA 540  
Db 367 AAGCAGTTAGACGGTGTGTCGGCAACTCCCAACAGACTTAGGTTTCTGTTGAGA 308  
Qy 541 GGGTGGACTCAGAGACAGGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT 600  
Db 307 TGGGGNACTCAGAGACAGGACTAGCTGGATYCCYAGGCTGACTAAGAAATCCCTAAGCCT 248  
Qy 601 ANCTGGGAAGGTGACCGCATCTTAAATGTTAAGATGGGGCTTGGCACTTAGCTCACACCG 660
```

Db 247 AGCTGGGAAGGTGACCATCCACCTTTAAACAGGGGCTTGCACTAGCTCACACTG 188

Qy 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC 720

Db 187 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC 128

Qy 721 AATCATCTATTGCTCAGAGACACAGCGGGAAGGACAGGATGGGATATAAATCAGGCA 780

Db 127 RATTATTTATTGCTCAGAGACACAGGAGGACGATATCGGATATAAATCAGGCA 68

Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATGATGGAGCTCTGTTTT 840

Db 67 TTCGAGCCGGCAACAGAAACCCYTTT-TTTCYCYCTTTTGTGTTGTTTNTT 9

Qy 841 CACTCTAT 848

Db 8 NTTATTT 1

RESULT 10
BX380176/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX380176 966 bp mRNA linear EST 28-APR-2004
BX380176 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1044YK06 3-PRIME, mRNA sequence.

BX380176
BX380176.2 GI:46834414
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 966)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30460243.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1044BF03NP1&c=4215.r.
Location/Qualifiers

FEATURES
Source
1..966
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1044YK06"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was
normalized."

ORIGIN
Query Match 45.9%; Score 610; DB 5; Length 966;
Best Local Similarity 82.5%; Pred. No. 2,2e-173;
Matches 693; Conservative 24; Mismatches 110; Indels 13; Gaps 2;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTATTTT 60
Db 830 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 771
Qy 61 AGGGGAAGATGCTGTAGTATGTTTAACTCAATCTGAATCATCTACGAGAAAGTTAAAGA 120
Db 770 AGGGGAAGATGCTGTATTTATTTATTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGA 711

Qy 121 AATTGAGATCGAATAATAATGTAGACGAGGAGACCTTCAAAAACATGCACTGCGGSCCT 180

Db 710 AATTGAGATCGAATAACAGTAGACGAGGAGCTTCGAAACACTGGACCTCGGSCCT 651

Qy 181 CCTCAGCAATGAGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATATATT 240

Db 650 CCTCAGCAATGAGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATATATT 591

Qy 241 TTTACTCTCTTTTGGACCCCTGTATCTTCAACTTCCCTTTTAAAGTTTGTCTCTCCAGAA 300

Db 590 GCTACTCTCTTTTGGACCCCTGTATCTTAACTTCCCTTTTAACTTTGTCTCTCCAGAA 531

Qy 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAATGGAAACCCAGATGCAATGCAATGATA 360

Db 530 CGAAGCTGTAAACTA-----CAATGGAGCCCGAGTGCAGTCCARGACTAG 483

Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420

Db 482 GDTCTACCGGCGCCCTGTCGACCGGCTGTGTCGCGCCCGMTCTGATGTTAATGACATCA 423

Qy 421 AGTCACCCCTCCGAGGAAATCTCACTGCACAAACCCCTACTACACTCCCAATTCAGTAGG 480

Db 422 AGGCACCCCTCTCTGGGCAATCTCAGCTGCACACCCCTCTACTACGCCCAATTCGCGGG 363

Qy 481 AAGCAGTTAGAGCAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540

Db 362 AAGCAGTTGGGGGCTGCTCGGCCAACCTCCCGACGCACTTTRGGTTTCTGTTGAGA 303

Qy 541 GGGTGGACTGAGACAGGACTAGTGGATTTCTAGGCTGACTAAGAATCCCAAGCCT 600

Db 302 TGGGGACTGAGACAGGACTAGTGGATTTCTAGGCTGACTAAGAATCCCTTARGCT 243

Qy 601 ANCTGGGAAGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660

Db 242 RGCTGGGAAGTGACCAATCCACCTTTAAACMGGGGCTTGCCACTTSGCTCACCCTG 183

Qy 661 ACCAATCAGAGAGCTCACTAAATGTCTAATCAGCAAAAAACAGGAGTAAAGCAATAGC- 719

Db 182 CCCAATCGGGGCTCACTAAATGTCTAATTCGCGCCCGCCSSSSSVSSSSGCGCTCGCG 123

Qy 720 CAATCATCTATTGCTCAGAGCAGCAGGGAAGGACAGGATTCGGATATATAACTCAGGC 779

Db 122 CMSVMTCTATTGCTGAGACACAGCMGGAGGACATGATCGGATATAAACCAGMT 63

Qy 780 ATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATGATGAGGAGCTCTGTTT 839

Db 62 CTTCGCGCGCCCGCGCCCGCCCTTTGGGTCCCTCCCTTTTCGTTNNTGGAGCTCTGT 3

RESULT 11
BX357208/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX357208 758 bp mRNA linear EST 08-APR-2004
BX357208 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1022YJ18 3-PRIME, mRNA sequence.

BX357208
BX357208.2 GI:46305595
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 758)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30376125.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?s=CS0DI0222DE09NP1&c=4215.r>.

FEATURES
SOURCE

```

1 758
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/class="CS01022xJ18"
/class_type="PLACENTA COT 25-NORMALIZED"
/class_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	45.8%	Score 508.2	DB 5	Length 758
Best Local Similarity	88.1%	Pred. No. 7.2e-173		
Matches	665	Conservative	9	Mismatches 69
			Indels	12
			Gaps	1
QY	75	GTTAGTATGTTAATCAATCTGGAAATCAATCTAGAGAAAGTTAAAGAAATTTGAGATCGAA	134	
DB	758	GTATTATTGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAA	699	
QY	135	TATAATCTAGACGACGAGCACTTCAAAACACTGCACCTCTGGGGCTCTCTCAGACCAATGGA	194	
DB	698	TACAACGTAGACGACGAGAGGTTTCAAACACTGAGACCTCTGGGGCTCTCTCAGACCAATGGA	639	
QY	195	TGCCCTCGACTCTCCCTTCTTAGGACCTCTAGCAGCTATATAATTTTTACTCTCTTTG	254	
DB	638	TGCCCTCGATTCTCCCTTCTTAGGACCTCTAGCAGCTATATAATTTGCTACTCTCTTTG	579	
QY	255	GACCTCTATCTTCAACTTCCTTTTAAAGTTTGTCTTTCAGAAATTTGAAGCTGTAAAGC	314	
DB	578	GACCTCTGATCTTTAAACCTCTTGTTAACCTTTGTCTCTTCAGAATCGAAGCTGTAAAC	519	
QY	315	TACAAATAGTCTTCAAAATGGAACCCAGATGCACTCCATCACTTAAATCTTACCGTGAC	374	
DB	518	TA-----CAATGGAGCCCAAGATGCACTCCAGCTTAAGATCTACCGCAGC	471	
QY	375	CCCTGGACCGGCTGTAGACTATGTCTCTGATGTTAATGACATTTGAAGCTCACCCCTCCG	434	
DB	470	CCCTGGACCGGCTGTAGCCAGATCTGATGTTAATGACATCAAGAGCAACCCCTCTG	411	
QY	435	AGGAAATCTCAACTTGCACAAACCCCTACTACACTCCAAATTCAGTAGGAAGCAGTTAGACA	494	
DB	410	AGGAAATCTCAGTGTCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGG	351	
QY	495	GTTGTACGCCAACCTCCCCAAACAGTACTTGGTTTTCTGTGTAGAGGGTGGAGTGAAG	554	
DB	350	GTCGTGGGCCAACCTCCCCAAACAGCAGTACTTGGTTTTCTGTGTAGATGGGGACTGAGAG	291	
QY	555	ACAGACTAGCTGGATTTCTTAGGTGACTAAGAAATCCNAAAGCCTACTCTGGGAAGTTGA	614	
DB	290	ACAGACTAGCTGGATTTCTTAGGTGACTAAGAAATCCCTAAGCCTAGCTGTGGGAAGTTGA	231	
QY	615	CCGATCCATCTTTAAACATGGGGCTTCCAACTTAGCTCACACCCGACCAATCAGAGAGC	674	
DB	230	CCATCCACCTTTAAACACGGGGCTTCCAACTTAGCTCACACCTGACCAATCAGAGAGC	171	
QY	675	TCACTAATATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGGCAATCATCTATTGCC	734	
DB	170	TCACTAATATGCTAATTAGGCAAAACAGGAGGTAAAGAAATGAGCAATCATCTATTSC	111	
QY	735	TGAGAGCACAGCGGGAAGGACAAAGGATTTGGGATATAAACTCAGGCAATTTCAAGCCAGCAAC	794	
DB	110	YSAGAGCACAGCAGAGGGAACATGATCGGGATATAAACCCMWSSYCCCCCGCCGCGCCAC	51	
QY	795	AGCAACCCCTTTTGGTCCCCCTCCCAATTTGATGGG	829	

50 GGCMACCCCCYCTGGGTCCCTCCCTTGTATGGG 16

RESULT 12	ACCESSION	REFERENCE
BI087886	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

BI087886 771 bp mRNA linear EST 20-JUN-2001
602852690F1 NIH_MGC_10 Homo sapiens CDNA clone IMAGE:4993894 5',
mRNA sequence.

ACCESSION	BI087886
VERSION	BI087886.1
	GI:14506216

KEYWORDS	EST.
SOURCE	Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 771)

NIH-MGC <http://mgc.nci.nih.gov/>.
Mammalian Gene Collection (MGC)

National Institute
 Published 1988Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Contact: Robert C. Anderson, MD
Email: ccanpb-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cdNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the TMCCE Consortium/INT at:

Found through the I.M.A.G.E. Consortium/ BENE at
<http://image.llnl.gov>

ncsp://image:tmr.gov
Plate: I.TAM11015 row: d column: 23

High quality sequence stop: 762.

 Location/Qualifiers |

source

/organism="Homo sapiens"

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/seq_id="IMACE.1993994"
```

```
/clone="IMAGE:4993034"
/cell_line="MGC36"
```

```
/cell_line= MOC30
/lab_host="DH10B"
```

```
/clone lib="NIH MGC 10"
```

/note="Organ: c̄ervix; Ve

Site_2: SalI; Cloned un...

Average insert size 1.5

22

Query Match	44.7%	Score 593.4	DB 4	Length 771
Best Local Similarity	89.2%	Pred. No. 2.3e-168		
Matches 688	Conservative 0	Mismatches 68	Indels 15	Gaps 4
QY	107	GAGAAAGTTAAAGAAATTTGAGATCGAATATAATGTAGACAGAGGAGCCTTCAAAACACT	166	
DB	1	GAGAAAGTTAAAGAAATTCGAGATCGAATACACGTAGACAGAGGAGCTTCGAAACACT	60	
QY	167	GCACCTGGGGCCTCCTCAGCCAATGGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTA	226	
DB	61	GGACCTGGGGCCTCCTCAGCCAATGGATGCCCTGGATTCTCCCTCTTTAGGACCTCTA	120	
QY	227	GCAGCTATAATATTTTTACTCCTCTTTGGACCTGTATCTTTCAACTCTCCTTGTTAAGTTT	286	
DB	121	GCAGCTATAATATTGCTACTCCTCTTTGGACCTGTATCTTTAACTCCTCTTGTTAACATT	180	
QY	287	GTCCTCTCCAGAAATTTGAAGCTGTAAAGCTACAAATAGTTCCTTCAATCGMAACCCAGATG	346	
DB	181	GTCCTCTCCAGAAATCGAAGCTGTAAAGCTA	228	
QY	347	CAGTCCATGACTAAATCTACCGCTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGAT	406	
DB	229	CAGTCCAGACTAAGATCTACCGCAGACCCCTGGACCGGCTGCTAGGCCAGATCTGAT	288	
QY	407	GTTAATGACATTGAAGTCCACCCCTCCCGAGGAAATCTCAA	466	
DB	289	GTTAATGACATCAAGGCACCCCTCTTGAGGAAATCTCAGGTGACAACTCTACTAGGC	348	


```
QY 991 GTGTCCACTGTCTCTGATCAGCGAGGTACCCATTGGCACTCCCGATCAGGCTAAAGG 1050
Db 1465 -----GATCAGCGAGCGCCCATTTGCTCTCTGATCGGCTAAAGG 1507
QY 1051 CTTGCCATTGTCCTGATGCTTAAGTGTCTGGTGTTCCTTAATAGAACTGAACACTGG 1110
Db 1508 CTTGCCCTTGTTCCTGCGAGGCTAAGTGCACAGTTTGGTCTTAATAGCTAAACACTAG 1567
QY 1111 TCACCTGGGTTCATGTTCTCTTCCATGACCCACGCTTCTAATAGAGCTATAAACACTCA 1170
Db 1568 TCACCTGGGTTCACAGGATCTCTTCCATGACCCACAGCTTCTAATAGAGCTATAAACACTCG 1627
QY 1171 CGCATGGCCCAAGATTCCATTCCTTGGTATCTGTGAGGCCCAAGAACCCAGGTACAGAGA 1230
Db 1628 CCACTTGGCCCAAGATTCCATTCCTTGGTATCCATGAGGCCCAAGAACCCAGGTACAGAGA 1687
QY 1231 ANGTGAGGTTGCCACATTTGGGAAGTGGCCACATGCCATTTGTAGCGGCCCAACAC 1290
Db 1688 ACACAGAGCTTGCCACCATCTCGGAAGTGGCCCGTCCACATCTTGGGAAGCGACCTGCCAC 1747
QY 1291 CATCTTGGGAGCTGTGGGAGCAAGGATCCCC 1321
Db 1748 CATCTTGGGAACCTTGGGAGCAAGGACCCCC 1778
```

```
RESULT 14
CN272394 719 bp mRNA linear EST 16-MAY-2004
LOCUS 170060005410 GRN_PNEU Homo sapiens cDNA 5', mRNA sequence.
CN272394
VERSION CN272394.1 GI:47288808
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 719)
Li,Y., Xu,C., Fang,R., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Brandenberger,R., Wei,H., Zhang,S., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,B.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 719 Std Error: 0.00.
Location/Qualifiers
1. 719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/mol_type="mRNA"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN_PNEU"
/notes="oligo dt primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."
```

```
FEATURES
source
Query Match 42.2%; Score 561; DB 7; Length 719;
Best Local Similarity 85.4%; pred. No. 1.6e-158;
Matches 657; Conservative 0; Mismatches 62; Indels 50; Gaps 1;
QY 95 GGAATCATTTCTGAGAAAGTTAAAGAAATTTGAGATCGAATATATGATAGCAGGAGGAC 154
Db 1 GGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAACTAGCAGAGGAG 60
```

```
QY 155 CTTAAAAACACTGCACCCCTGGGGCTCTCTCAGCCAAATGATGCCCTGGACTCTCCCTTC 214
Db 61 CTTGAAAACACTGGAGCCCTGGGGCTCTCTCAGCCAAATGATGCCCTGGACTCTCCCTTC 120
QY 215 TTAGACCTCTAGCAGCTATAATATTTTACTCTCTTTGGACCTGTATCTTCAACTTC 274
Db 131 TTAGACCTCTAGCAGCTATAATATTTTACTCTCTTTGGACCTGTATCTTCAACTTC 180
QY 275 CTTGTTAAAGTTTGTCTCTCTCCAGAAATGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATG 334
Db 181 CTTGTT-----AAATG 190
QY 335 GAAACCCAGATGACGATCCATGATTAATCTACCGTGGACCCCTGGACCGGCTGTCTAGA 394
Db 191 GAGCCCAAGATGACGATCCCAAGACTAAGATCTACCCAGACCCCTGGACCGGCTGTCTAGC 250
QY 395 CTATCTCTGATGTTAATGACATTCAGATTCAAGTCAACCCCTCCCGAGGAATCTCAACTGCACAA 454
Db 251 CCACGATCTGATGTTAATGACATCAAGGACCCCTCTCTGAGGAATCTCAGCTGCACAA 310
QY 455 CCCCTACTACACTCCCAATTCAGTAGGAAGCAGTTAGAGCAGTTGTGAGCAACCTCCCCA 514
Db 311 CTTCTACTAGCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCTGCGGCAACCTCCCCA 370
QY 515 ACAGTACTTGGGTTTTCTCTGTTGAGGGTGGAGTGGAGCAGAGCAGTCTGATGATTTCC 574
Db 371 ACAGCATTAGGTTTTCTCTGTTGAGATGGGGAGTGGAGCAGAGCAGTCTGATGATTTCC 430
QY 575 TAGGCTGACTAAGATCCNAGCCCTANCTGGGAAGTGCACCGCATCTTAAACAT 634
Db 431 TAGGCTGACTAAGATCCCAAGCCCTAGCTGGGAAGTGCACCATCTCACTTTAAACAC 490
QY 635 GGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAAATGCTAATCAGG 694
Db 491 GGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAATGCTAATAGG 550
QY 695 CAAAACAGAGGTAAGCAATAGCAATCATCTATTGCTGAGAGCAGCAGCGGGAAGGA 754
Db 551 CAAAGACAGAGGTAAGAAATAGCAATCATCTATTGCTGAGAGCAGCAGCAGGAGGA 610
QY 755 CAAAGATTGGGATATAAATCTCAGGATTCAGCCAGCAACAGCAACCCCTTTGGGTCCC 814
Db 611 CAATGATCGGATATAAACCAGCTCTCGAGCGGCAACGGCAACCCCTTTGGGTCCC 670
QY 815 CTCCCATTTGATGGGAGCTCTGTTTTCACCTCTATTTCATCTATTAAAT 863
Db 671 CTCCCTTTGATGGGAGCTCTGTTTTCATGCTATTTCATCTATTAAAT 719
```

```
RESULT 15
BX459153 900 bp mRNA linear EST 05-MAY-2004
LOCUS BX459153 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX459153
VERSION BX459153.2 GI:47051796
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 900)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31021096.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
```


was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 4215.1

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna78=CS0DE013BE10NPI&c=4215.1.

FEATURES

Location/Qualifiers

1..900

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DE013Y120"

/issue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 40.9%; Score 544; DB 5; Length 900;

Best Local Similarity 77.5%; Pred. No. 2.5e-153;

Matches 631; Conservative 49; Mismatches 120; Indels 14; Gaps 3;

QY 1 TCAAAATCGAAGCTTTAGACTTCTAACCAGCCCAAGAGGGGAACTGTTATTTT 60

DB 802 TCANAATCGAAGCTTTAGACTTCTAACCAGCCCAAGAGGGGAACTGTTATTTT 743

QY 61 AGGGGAAGAAATGCTTTAGTATGTTAATCAATCTGGAATCAATTAAGAGGTTAAAGA 120

DB 742 AGGGGAAGAAATGCTTTAGTATGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAAGA 683

QY 121 AATTTGAGATCGAATATAATATAGAGCAGAGACCTTCAAAACACTGACCCCTGGGGCT 180

DB 682 AATTCGAGATCGAATACAACTAGAGCAGAGGAGTTTCGAAACACTGGACCCCTGGGGCT 623

QY 181 CCTCAGCAATGGATGCTGACCTCCCTCTTAGGACCTCTAGAGCTTATAATATT 240

DB 622 CCTCAGCAATGGATGCTGACCTCCCTCTTAGGACCTCTAGAGCTTATAATATT 563

QY 241 TTACTCTCTTTGGACCTGATCTTCAACTTCTTTGTTAAGTTGTTCTTCCAGAT 300

DB 562 GCTACTCTCTTTGGACCTGATCTTCAACTTCTTTGTTAAGTTGTTCTTCCAGAT 505

QY 301 TGAAGCTGAAGCTTCAAAATAGTTCTTCAATAGGAACCCGATGTCAGTCCATCACTAA 360

DB 504 CGAAGCTGTAACCA-----CAAAATGGAGCCCAAGATGCAAGCCCAACCAAA 457

QY 361 AATCTACCTGGACCTCGACCGCTCTAGACTATCTCTGATGTTAATGACATTGA 420

DB 456 SACTTACCGAGACCTCTGACCGCGCTGACCGCGCTGACCGCGCTGACCGCGCTGAC 397

QY 421 AGTCACCTCTCCGAGGAATCTCAACTGCAACCCCTTACTACTCACTCAATTCAGTAGG 480

DB 396 ASSCACCCCCGAGGAATCCAGCGCAGCAACCTCTACTACGCCCCCAATTCAGCAG 337

QY 481 AAGCAGTTAGAGCTTGTGAGCAACCTCCCAACAGTCTGGGTTTCTGTTGAGA 540

DB 336 AAGCAGTTAGAGCTTGTGAGCAACCTCCCAACAGTCTGGGTTTCTGTTGAGA 277

QY 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600

DB 276 TGGGCACTTSGAGASACGACTAGCTTSSATCCMCCMCCMCCMCCMCCMCCMCCMCC 217

QY 601 ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGCACTTAGCTCACACCCG 660

DB 216 MGCTGGGAAGGTGACCAACCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCC 157

QY 661 ACCAATCAGAGACTCACTAAATGCTTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 720

DB 156 SCCAATCAGAGCGGSGGAAGGCTTCTAGGCAAGAGCAGAGCGAGGRRATMGCC 97

QY 721 AATCATCTATTGCTTGAGAGCAGCGGGGAAGGAAGGATTGGGGATATAAACTCAGGCA 780

DB 96 MVCCATSCATGCTCTGAGAGCACRGCCAGGAGGACVATGCGGGAGRTAAASCCAAGTC 37

QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGTCCC 814

DB 36 BCCGCGCGGCTACGGCKCCCCCTTKTGGTCCC 3

RESULT 16

BX953822/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1002)

Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,

Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and

Wiemann,S.

EST (Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., et al.)

Unpublished (2003)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by EMBL (European Molecular Biology Laboratories,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No 5' sequence available.

This clone (DKFzp781M064) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..1002

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="DKFzp781M064"

/dev stage="adult"

/lab_host="DH10B"

/clone_lib="781 (synonym: hlcc4)"

/notes="vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;

cDNA-collection"

ORIGIN

Query Match 37.7%; Score 500.6; DB 5; Length 1002;

Best Local Similarity 87.6%; Pred. No. 4e-140;

Matches 641; Conservative 0; Mismatches 72; Indels 19; Gaps 8;

QY 549 TGAGAGCAGCAGCTAGCTGGATTTCTAGGCTGATCCAGAAATCCNAAGCCTTCTGGGA 608

DB 922 TGAGAGCAGAACTAGCTGGATTTCTAGGCAACTAAGAAATCCGCTAGCTGGGA 863

QY 609 AGGTGACCGCATCATCTTTAAACATGGGGCTTCGAACCTTGGCTCACACCCGCAATCA 668

DB 862 A-GTGACCGCATCCACCTTTAAACACGGGGGCTTGAACCTTGGCTCACACCCGAGAG 808

QY 669 GAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCT 728

DB 807 -----CTCCTTAAATGCTAATTAGGCAAAACAGGAGGTAAAGAAATAGCCGATCATCT 753

QY 729 ATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCATTTCAAGCC 788

DB 752 ATTGCTGAGAGCAGCAGC-GGAAGGACAAGGATCGGGATATAAAACAGGCATTCGAGCC 694

QY 789 AGCAACACCAACCCCTTTGGGTCCTCCCTCCCTGATGGAGCTCTGTTTTCTACTCTAT 848

/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-120G11.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 36.2%; Score 480.8; DB 9; Length 679;
Best Local Similarity 86.7%; Pred. No. 3.7e-134;
Matches 552; Conservative 0; Mismatches 83; Indels 2; Gaps 2;

QY 673 GTCCTAAATGCTAAATCAGGCAAAAACAGAGGTAAGCAATAGCCATCATCTATTG 732
DB 678 GCTCGTAAATGCTAAATCAGGCAAAAACAGAGGTAAGCAATAGCCATCATCTATTG 619
QY 733 CTTGAGAGCAGGCGGAGGAGGAGGATTTGGGATATAAATCAGGCAATTCAGGCGAGCA 792
DB 618 CTTGAGAGCAGGCGGAGGAGGAGGATTTGGGATATAAATCAGGCGAGCA 559
QY 793 AC-AGCAACCCCTTTGGGTCCTCCCATTTGATGGGAGCTCT-CTTTTCACTCTATTT 850
DB 558 AGGGGAGGCGGCTTTGGGTCCTCCCTCTTTGATGGGAGCTCTGTTTCACTCTATTT 499
QY 851 CACTCTATTAAATCATGCAATGCACTCTTTGGTCCGCTGTTTATGGCTCAAGCTGA 910
DB 498 CACTATATTAGATCTTTGCACTGCACTCTTTGGTCCGCTGTTTATGGCTCAAGCTGA 439
QY 911 GCTTTTGGTCCGCTCACTGCTGTTTGGGTCCTCCATGCTGCTGCTGCTGCTGCTGCT 970
DB 438 GCTTTTGGTCCGCTCACTGCTGTTTGGGTCCTCCATGCTGCTGCTGCTGCTGCTGCT 379
QY 971 ATCCCTTTGGTCCGCTCACTGCTGTTTGGGTCCTCCATGCTGCTGCTGCTGCTGCTGCT 1030
DB 378 ATCCCTTTGGTCCGCTCACTGCTGTTTGGGTCCTCCATGCTGCTGCTGCTGCTGCTGCT 319
QY 1031 ACTCCGATCAGGCTAAAGGCTTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090
DB 318 GCTCCGATCAGGCTAAAGGCTTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
QY 1091 CTAAATGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
DB 258 GTAATGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199
QY 1151 TAATAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
DB 198 TAATAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 139
QY 1211 CAAGAACCCGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1270
DB 138 CAAGAACCCGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 79
QY 1271 TTTTGTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1307
DB 78 TTTTGTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 42

RESULT 18
BM918330/c
LOCUS
DEFINITION
AGENCY: NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5747240
5', mRNA sequence.
BM918330
BM918330.1 GI:19368709
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 999)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

REFERENCE
AUTHORS
TITLE
JOURNAL

693 AGCAATGGCAACCCCTTTGGGTCCTCCCTCGATGGAGCTCTGTTTCACTCTAT 634
QY 849 TTCACCTATTAATCAATGCAATGCACTCTTCTGCTCGGTGTTTATGCTCAAGCT 908
DB 633 TTCACACTATTAATCAATGCAATGCACTCTTCTGCTCGGTGTTTATGCTCAAGCT 574
QY 909 GAGCTTTTGGGTCGCAATGCACTCTGTT-TGCCACCGTCAACAGCCGCTGCTGACT 967
DB 573 GAGCTTTTGGGTCGCAATGCACTCTGTTTGGGTCGCTGCTGCTGCTGCTGCTGACT 514
QY 968 TCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
DB 513 TACATTTCTCCGATCCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 454
QY 1027 TGGCACTCCGATCAGGCTAAAGGCTTGCCA-TTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1085
DB 453 TGGCGCTCCAGATCGGGCTAAAGGCTTGCCA-TTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 394
QY 1086 TTGCTCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145
DB 393 TTGCTCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334
QY 1146 GCTTTCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
DB 333 ACTTCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 278
QY 1206 GAGGCAAGAACCCGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1264
DB 277 GAGGCAAGAACCCGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 218
QY 1265 CTGCCATTTGG 1276
DB 217 CGACCATCTTTG 206

RESULT 17
AG113694/c
LOCUS
DEFINITION
Pan troglodytes DNA, clone: PTB-120G11.R, genomic survey sequence.
AG113694
AG113694.1 GI:16734213
VERSION
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 679)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@gsic.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .679
/organism="Pan troglodytes"

COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12773 row: b column: 09 High quality sequence stop: 685. Location/Qualifiers 1..999 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5747240" /lab_host="DH10B" /clone_lib="NIH_MGC_120" /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."			
	FEATURES source			
	ORIGIN			
	Query Match 36.2%; Score 480.8; DB 5; Length 999; Best Local Similarity 83.1%; Pred. No. 4.1e-134; Indels 16; Gaps 6; Matches 621; Conservative 0; Mismatches 110			
QY	549	TGAGACACGAGGACTAGCT-GGATTTTCTAGGCTGACTAAGAAATCCCAAGCCTACTGCGG	607	
Db	886	TGAGACACGAGGACTAGCTGGGATTTCTAAGCGGACTAAGAAATCCCTAGGCTAGTTAGG	827	
QY	608	AGGTGACCGATCCATCTTTAA-CATGGGCTTGCACTTAGCTCACACCGACCAAT	666	
Db	826	AAGGTGACCGCATCCACCTTAAACCATGGGCTTGCAACTTAGCTCACACCGACCAAT	767	
QY	667	C-----AGAGACTCAGTAAATGCTAATCAGGC-AAAAACAGGAGGTAAGCAAT	716	
Db	766	CAGGTAGTAAGAGAGCTACTAAATGCCATCAGACAAAACAGGAGGTAAGAACT	707	
QY	717	AGCAATCATCTATTGCTTGAGAGCACAGCGGAAGGACAGGATGGATATAACTCA	776	
Db	706	AACCAATCATCTATTGCTTGAGAGCACAGTGGGAGGACAATGATCAGGATATAAGCCA	647	
QY	777	GGCAATTCAGCCAGCAACAGCACCCCTTTGGGTCCTCCCATTTGATGGAGCTCG	836	
Db	646	GGCAATTCAGCCAGCAACAGCGCTACCCCTCTTTGGCTCCCTCCCTTTGATGGAGCTGG	587	
QY	837	TTTTCACTCTATTTCCTACTCTATTAAATCATGCAACT---GCACTCTTCTGGTCCGCTTTT	893	
Db	586	TTTTCACTCTATTAAATCTTGTAAGTCATCTCTTGTGTCGCTGTTGTACCGCTGTTT	527	
QY	894	TTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCAGCTCACAG	953	
Db	526	GTTTACCGCTCAAGCTTAAGCTTTGCTGCTGTTTGCACCACTGCTGTTTGGCGAGTCGAG	467	
QY	954	ACCGCTGCTGACTTCCATCCCTTTGGATCCACAGAGTGTCCACTGTGCTCTGATCCA	1013	
Db	466	ACCGCGCGCTGACTTCCACCCCTCCGGATCTGCAGGGGTGCTGCTGTCTCTGATCCA	407	
QY	1014	GCAGGTGACCAATTCGCACTCCCGATCAGGCTAAAGCTTGCATTGCTTCATGGCT	1073	
Db	406	GCAGGCGCCCACTGCTGCTCCCGATCGGGCTGAAGGCTTGCCGCTTGTTCCTGCACGGC-	348	
QY	1074	AAGTGCCTGGTTTGCTCTTAATAGAACTGAACACTGGTCACTGGTTCATGGTCTCTTT	1133	
Db	347	AAGTGCCTGGTTTGCCTCTTAATTAGATGAACACTAGTCTGCTGGGTTCTACGGTCTCTTT	288	
QY	1134	CCATGACCCACGCGTTCTTAATAGAGCTATAAACACTCACCGCATGGCCCAAGATTCCATTC	1193	
Db	287	CCATGACCCACGCTCTTCCATAGAGCTATAAACAATTCACCGCATGGCCCAAGATTCCATTC	228	
QY	1194	CTTGGTATCTGTGAGGCCAAGAACCCAGGCTCAGAGAAAGTGAAGGCTTGCCACCATTTGG	1253	
Db	227	CTTGGAAATCGTGGGGCCAAGAACCCAGGCTCAGAGAACGAGAGGCTTGTGCCATCTTG	168	
QY	1254	GAAGTGGCCCACTGCCATTTTGGTAGC	1280	
Db	167	GAAGCAACCCGCCACCATGTTGGAGC	141	
RESULT 19				
AG134524	AG134524 683 bp DNA linear GSS 04-NOV-2001			
LOCUS	Pan troglodytes DNA, clone: PTB-147110.R, genomic survey sequence.			
DEFINITION	AG134524			
ACCESSION	AG134524.1 GI:16664202			
VERSION	GSS.			
KEYWORDS	Pan troglodytes (chimpanzee)			
SOURCE	Pan troglodytes			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.			
REFERENCE	1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. BAC end sequences of Library PTB Unpublished			
AUTHORS	2 (bases 1 to 683) Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. Direct Submission			
TITLE	Submitted (02-AUG-2001) Aaso Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimps@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)			
JOURNAL	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.			
COMMENT	PRIMERS Sequencing: M13Rev LIBRARY Vector : pKS145 R.Site 1 : SacI R.Site 2 : SacI Location/Qualifiers 1..683 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" /clone="PTB-147110.R" /sex="male" /cell_type="lymphoblast" /clone_lib="PTB Chimpanzee Male BAC Library"			
FEATURES	source			
ORIGIN	Query Match 35.8%; Score 475.2; DB 9; Length 683; Best Local Similarity 89.6%; Pred. No. 1.8e-132; Matches 554; Conservative 0; Mismatches 60; Indels 4; Gaps 4;			
QY	666	TCAGAGAGCTCACTAAAATGCTAATTCAGGCAAAAACAGGAGTAAGCAATAGCCAAATCA	725	
Db	58	TCGNAGAGCTCACTAAAATGCTAATTCAGGCAAAAACAGGAGTAAGCAATAGCCAAATCA	117	
QY	726	TCTATTGCTTGAGACACAGCGGAGGACAGAGGATTCGGATATAAACTCAGGCAATTCAA	785	
Db	118	TCTATTGCTTGAGACACAGTGGGAGGACAGAGGATTCGGATATAAAACCCAGGCAATTC	177	
QY	786	GCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTTCATCTC	845	

Db 178 GCAGCAACGGCAACCCCTTTGGT-CCTCCCTTGTATGGAGCTCTGTTTTCACATC 236
 QY 846 TATTTCACTCTATTAATCATGAACACTGCACTCTCTGTCGCTGTTTATATGCTCAA 905
 Db 237 TATTTCACTCTATTAATCATGAACACTGCACTCTCTGTCGCTGTTTATATGCTCAA 296
 QY 906 GCTGAGCTTTTGTTCGCCATCCACCACTGCTG-TTGGCCACCGTCAACAGCCGCTGCTG 964
 Db 297 GCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCGCTGCAACAGCCGCTG 356
 QY 965 ACTTCCATCCCTTTGGATCCAGCAGAGTGTCACTGCTGCTGATCCAGCAGGTAACCC 1024
 Db 357 ACTTCCATCCCTCCAGATCCAGCAGGCTGCTGCTGCTGATCCAGCAGGTAACCC 416
 QY 1025 ATGGCACTCCCATGAGCTTAAGCTTGCCTGCTGCTGCTGATCCAGCAGGTAACCC 1084
 Db 417 ATGGCGCTCTGATCGGCTAAAGCTTGCCTGCTGCTGCTGATCCAGCAGGTAACCC 476
 QY 1085 TTTGTCCTTAATGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1144
 Db 477 TTTGTCCTTAATGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536
 QY 1145 GGTCTCTAATAGAGCTATTAACACTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1203
 Db 537 GACTTCTAATAGAGCTATTAACACTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
 QY 1204 GTAGGCGCAA-GAAGCCAGGTCAGAGANGTGAAGCTTGCCTGCTGCTGCTGCTGCTG 1262
 Db 597 ATGAGGCGCAATGAAGCCAGGTCAGAGANGTGAAGCTTGCCTGCTGCTGCTGCTGCTG 656
 QY 1263 CACTGCGCATTTTGGTAGC 1280
 Db 657 CACCACATCTGGGAAC 674

RESULT 20
 AG049953 669 bp DNA linear GSS 02-NOV-2001
 LOCUS Pan troglodytes DNA, clone: PTB-030D19.F, genomic survey sequence.
 AG049953
 AG049953.1 GI:16596845
 VERSION
 GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 1
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB
 Unpublished
 2 (bases 1 to 669)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimpanzee@gsic.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.
 PRIMERS
 Sequencing: -21M13
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.
 Location/Qualifiers
 1. .669
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"

/db_xref="taxon:9598"
 /clone="PTB-030D19.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

Query Match 35.7%; Score 474; DB 9; Length 669;
 Best Local Similarity 86.8%; Pred. No. 4.2e-132;
 Matches 548; Conservative 0; Mismatches 71; Indels 12; Gaps 2;

ORIGIN

QY 669 GAGAGCTCACTAAATGCTATATCAGGCAAAAACAGAGAGTAAGCAATAGCAATCATCT 728
 Db 49 GAGAGCTCACTAAATGCTATATAGGCAAAAACAGAGAGTAAGCAATAGCAATCATCT 108
 QY 729 ATTGCTCTGAGACACAGCGGGAAGGACAGAGATGCGGATATAAACTCAGGCAATTCAGGC 788
 Db 109 ATTGCTCTGAGACACAGCGGGAAGGACAGATGATCGGATATAAACTCAGGCAATTCAGGC 168
 QY 789 AGCAACAGCAACCCCTTTGGGTCCCTCCCAATGATGAGGAGCTCTGTTTTCATCTAT 848
 Db 169 GGCACGGCTACTCTTTTGGGTCCCTCCCTCTGATGAGAGCTCTGT-----T 218
 QY 849 TTCACCTCTATTAATATGCAACTGCACTC--TTCGTGTCGCTGTTTATGCTGCTCAAG 906
 Db 219 TTCACCTCTATTAATATGCTGCAACTGCACTCTTTTATGTTTCCATGTTTGTACAGCTCGAG 278
 QY 907 CTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGCACCGTCCAGACCCGCTGCTGAC 966
 Db 279 GTGAGCTTTTGTATGCGGCTCCACCACTGCTGTTTGCCTGCGAGACCCCGCTGAC 338
 QY 967 TTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCAGAGTACCCAT 1026
 Db 339 TTCCATCCCTTTGGATCCAGCAGAGTGTCCCGTGTGCTCTGATCCAGCAGAGTACCCAT 398
 QY 1027 TGCCACTCCGATCAGGCTAAAGGCTTGCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
 Db 399 TGCTGCTCTGATCGGATTAAGGCTTGCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 458
 QY 1087 TGCTGCTTAATAGCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
 Db 459 CGTCTTAATAGCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 518
 QY 1147 CTTCTAATAGAGCTATAACACTCAGCAGTGGCCCAAGATTCATTCCTTTGGTATCTGTG 1206
 Db 519 CTTCTAATAGAGCTATAACACTCAGCAGTGGCCCAAGATTCATTCCTTTGGTATCTGTG 578
 QY 1207 AGGCCAAGAACCCAGGTCAGAGANGTGAAGCTTGCCTGCTGCTGCTGCTGCTGCTGCT 1266
 Db 579 AGGCCAAGAACCCAGGTCAGAGANGTGAAGCTTGCCTGCTGCTGCTGCTGCTGCTGCT 638
 QY 1267 GCCATTTTGGTAGCGGCCACCACTCTTG 1297
 Db 639 ACCATTTTGGTAGCGGCCACCACTCTTG 669

RESULT 21
 AG116971 681 bp DNA linear GSS 03-NOV-2001
 LOCUS Pan troglodytes DNA, clone: PTB-124K18.F, genomic survey sequence.
 AG116971
 AG116971.1 GI:16737490
 VERSION
 GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 1
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB
 Unpublished
 2 (bases 1 to 681)

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 681
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-124K18.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 35.6%; Score 473; DB 9; Length 681;
Best Local Similarity 88.8%; Pred. No. 8.6e-132;
Matches 538; Conservative 0; Mismatches 56; Indels 12; Gaps 2;
QY 671 GAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 730
DB 86 GAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 145
QY 731 TCCTTGAGAGCAGCGGGAAGGACAGAGGATGGGATATAACTCAGGCATTCAAGCCAG 790
DB 146 TGCTTGAGAGCAGCAGGAGGACCAATGATGGGATATAACCAGGATTCGAGCCAG 205
QY 791 CAACAGCAACCCCTTTGGGTCCTCCATTCATGATGGAGGCTCTGTTTCACTATTT 850
DB 206 CAAGGCTACCTCTTTGGGTCCTCCCTCTCTGATGGAGGCTCTGT-----TTT 255
QY 851 CACTCTATTAATCATGCAACTGCA--CTCTTCTGGTCCGCTGTTTATGGCTCAAGCT 908
DB 256 CACTCTATTAATCTTTGCAACTGCACTCTCTTCTGGTCCGCTGTTTATGGCTCAAGCT 315
QY 909 GAGCTTTTGTGCGCCATCCACACTGCTGTTTGCCACCGTCACAGACCGCTGCTGACTT 968
DB 316 GAGCTTTTGTGCGCCATCCACACTGCTGTTTGCCGCGCCGCGAGACCGCTGCTGACTT 375
QY 969 CCATCCCTTTGGATCCAGCAGAGTGTCACCTGCTCTCTGATCCAGCGAGGTACCATTTG 1028
DB 376 CCATCCCTCCAGATCTGGCAGGCTGTCACCTGCTCTCTGATCCAGCGAGGCGCCATTG 435
QY 1029 CCACTCCGATCAGGCTAAGGCTTGCCATGTTCTGTCGATGCTGCTAAGTGCCTGGGTTTG 1088
DB 436 CGCTCCCGATAGGGCTAAGGCTTGCCATGTTCTGTCGATGCTGCTAAGTGCCTGGGTTTG 495
QY 1089 TCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCCTTCCATGACCCAGGCT 1148
DB 496 TCTTAATAGAACTGAACACTAGTACCGGCTTCATGGTTCCTTCCGTCGACCCAGGCT 555
QY 1149 TCTTAATAGAACTGAACACTACCGGATGCGCCCAAGATTCATTCCTTGGTATCTGTGAG 1208
DB 556 TCTTAATAGAACTGAACACTACCGGATGCGCCCAAGATTCATTCCTTGGTATCTGTGAG 615
QY 1209 GCCAAGAACCCAGGTCAGAGANGTGAGGCTTGGCCACCATTTGGGAAGTGGCCCATCGC 1268
DB 616 GCCAAGAACCCAGGTCAGAGAACACAAAGGCTTGGCCACCATTTGGGAAGTGGCCCATCGC 675
QY 1269 CATTTT 1274

Db 676 CATTTT 681
|||||
AG096341 710 bp DNA linear GSS 03-NOV-2001
Pan troglodytes DNA, clone: PTB-097K23.F, genomic survey sequence.
AG096341
AG096341.1 GI:16716858
GSS.
Pan troglodytes (chimpanzee)
ORGANISM
SOURCE
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 710)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 710
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-097K23.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 35.3%; Score 469.6; DB 9; Length 710;
Best Local Similarity 85.1%; Pred. No. 9.4e-131;
Matches 561; Conservative 0; Mismatches 65; Indels 33; Gaps 2;
QY 671 GAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 730
DB 75 GAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 134
QY 731 TGCTTGAGAGCAGCAGCGGGAAGGACAGGATTTGGGATATAAATCAGGCATTCAAGCCAG 790
DB 135 TGCTTGAGAGCAGCAGTGGGATGGACATGACCAGGATATAAACCAGGCAATCCAGCCTG 194
QY 791 CAACAGCAACCCCTTTGGGTCCTCCATTCCTGATGGAGGCTCTGTTTCACTATTT 850
DB 195 CAACGGCAACCCCTTTGGTTCCTCTCTTTGATGGAGGCTCTGT-----TTT 244
QY 851 CACTCTATTAATCATGCAACTGCTCTCTTCTGGTCCGCTGTTTATGGCTCAAGCTGA 910
DB 245 CACTCTATGCAATCTGCAACTGCACTCTTCTGGTTCGTTGTTGATCGGCTCAAGCTGA 304
QY 911 GCTTTGTTGCGCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCCGTGTGCTTCC 970
DB 305 GCTTTGTTGCTCACCATCCACCACTGCTGTTTGGCCGCGTTGCGAGCCCATCGTGTCTCC 364

Fri Feb 25 16:26:29 2005

```

5', mRNA sequence.
ACCESSION BQ437925
VERSION BQ437925.1 GI:21177001
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 893)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13513 row: 1 column: 13
High quality sequence stop: 650.
FEATURES
    source
        1..893
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6161436"
            /tissue_type="melanotic melanoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 2 kb. Library constructed by Life
            Technologies."
ORIGIN
    Query Match 35.0%; Score 465.2; DB 5; Length 893;
    Best Local Similarity 80.4%; Pred. No. 2.2e-129;
    Matches 599; Conservative 0; Mismatches 101; Indels 45; Gaps 3;
QY 586 AGAATCCNAGGCTTANCTGGGAGGTGACCCATCCATCTTTAAACATGGGCTTGCAA 645
Db 1 AGAATCCCTAAGCTAGCTAGCTGGGAAATTTGACCACTGCTTAAACACGGGCTTGCAA 60
QY 646 CTTAGCTCACACCGGACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCA 696
Db 61 TTTAGCTCACACCGGACCAATCAGGTAGTAGTAAGAGAGCTCACTAAATGCTAATAGGGA 120
QY 697 AAAACAGGAGGTAAAGCAATAGCAATCATCTATTGCTGAGAGCAGCGGGAAGGACA 756
Db 121 AAAACAGGAGGTAAAGCAATAGCAATCATCTATTGCTGAGAGCAGCGGGAAGGACA 180
QY 757 AGAATGGGNTAATAACTCAGGATTCAGCCAGCAACAGCAACCCCTTTGGTCCCT 816
Db 181 ATGATCAGGATATAAACCAGGCAATTCAGCCAGCGGTGCTACCTCTTTGGTCCCT 240
QY 817 CCCATTGTATGGGAGCTGTGTTTTTCACTCTATTTCATCTATTAAATCATGCAACTGCAC 876
Db 241 CCCTTTGTATGGAAGCTCTGTTTTTCACTCTATTAAATCTTGAATTTGCA-----CAC 292
QY 877 TCTTCTGGTCCGTTTTTTATGGCTCAAGCTGAGCTTTTGTTCGGCATCCACCACTGCT 936
Db 293 TTTTCTGGTACGTGTGTGCACAGCTCAAGCTGAGCTTTGCTCACCCTCCACCACTGCT 352
QY 937 GTTTTGCACCGTTCACAGACCGCTGCTGACTTCCATCCCTTTGGATCCAGCAGGTGTC 996
Db 353 GTCTGCGCTGTTCACAGACCCACAGCTGACTTCCATCCCTCTGG----- 396
QY 997 ACTGTGCTCTGTATCCAGCGAGGTACCCATGTCATCTCCGATCAGGCTAAAGGCTTGCC 1056
Db 397 -----ATCCAGCGAGCGGCCCATTTGCTGCTCTCTGATCGGCTAAAGGCTTGCC 444

```

```

1057 ATTGTTCTCTGCATGGCTAAAGTGCCTGGTTGTCTTAATAGAACTGAACACTGCTACTG 1116
445 CTTGTTCTCTGCATGGCTAAAGTGCCTGGTTGTCTTAATAGAACTGAACACTGCTACTG 504
1117 GGTTCATGTTCTCTTCCATGACCCACGCTTTCTTAATAGAGCTATAAACAATCAGCAT 1176
505 GGTTCACGATTTCTTCCATGACCCACGCTTTCTTAATAGAGCTATAAACAATCAGCAT 564
1177 GGCCCAAGATTCATTCCTTGGTATCTGTGAGCCAAAGAACCCAGGTCTAGAGAAAGTGA 1236
565 GGCCCAAGATTCATTCCTTGGTATCTGTGAGCCAAAGAACCCAGGTCTAGAGAAACAA 624
1237 GGCTTGCACCAATTTGGGAAGTGCCTGCTGCTATTTGTTAGCGGCCACCACTCTTT 1296
625 GACTTGCACCAATCTCGGAAGTGCCTGCTGCTATTTGTTAGCGGCCACCACTCTTT 684
1297 GGAGAGCTGTGGGAGCAAGGATCCCC 1321
685 GGGNACTTGGGAGCAAGGACCCCC 709

RESULT 25
AG072852 Pan troglodytes DNA, clone: PTB-064H14.R, genomic survey sequence.
LOCUS AG072852
DEFINITION Pan troglodytes DNA, clone: PTB-064H14.R, genomic survey sequence.
ACCESSION AG072852
VERSION AG072852.1 GI:16624654
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 670)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Shuhiro-chou, Tsukuba, Ibaraki, Japan, URL: http://hgp.gsc.riken.go.jp/,
(E-mail: chimbe@gsc.riken.go.jp, Fax: 81-45-503-9170)
Tel: 81-45-503-9111.
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..670
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-064H14.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 34.3%; Score 456.2; DB 9; Length 670;
Best Local Similarity 86.4%; Pred. No. 1.1e-126;
Matches 529; Conservative 0; Mismatches 74; Indels 9; Gaps 2;
QY 669 GAGAGCTCACTAAATGCTTAATCAGGCAAAACAGGAGTAAGCAATAGCAATCATCT 728

```

```

Db 57 GAGAGCTCACTAAAAAGCTAATTAGGCAAAAAACAGGAGGTAAAGAAATAGCCAAATCATTT 116
QY 729 ATTGCTGTGAGACACAGCGGAGGACAGAGTGGGATATTAACCTACGACATTCAGGC 788
Db 117 ATTGCTGTGAGACACAGCGGAGGACAGATGATCGGATATTAACCCAGGACATTCAGGC 176
QY 789 AGCAACAGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTTCACCTCTAT 848
Db 177 AGCAAGGCTACCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTTTTTCACCTCTAT 236
QY 849 TTCACCTCTAATAATCATGCACTGCACTCTTCTGTCGCTGTTTATATGGCTCAAGCT 908
Db 237 TAAATCTTCAACTGC-----GCTCTCTTCTGCTGCGTTTGTACAGCTCAAGCT 288
QY 909 GAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCTCTGACTT 968
Db 289 GAG-TTTTGTGCTGTGTCACCACTGCTGTTTCCGCGTCGAGACCCGCTCTGAGTT 347
QY 969 CCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCGAGGTACCCATTG 1028
Db 348 CCATCCCTCGGATCCACAGGCTGTCCACTGTGCTCTGATCCAGCAAGGCGCCATTG 407
QY 1029 CCACTCCGATCAGGCTAAAGGCTTGCATTTCTGCAATGGCTAAGTCCCTGGGTTG 1088
Db 408 CCACTCTGATCGGCTAAAGGCTTGCATTTCTGCAAGGCTAAGTCCCTGGGTTG 467
QY 1089 TCCTAATAGACTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148
Db 468 TCCTAATAGACTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
QY 1149 TCTAATAGACTGATAACACTCAGCGATGCGCCCAAGATTCCATTCCTTGTGATCTGTGAG 1208
Db 528 TCTAATAGACTGATAACACTCAGCTGATGCGCCCAAGATTCCATTCCTTGTGATCTGTGAG 587
QY 1209 GCCAAGAACCCAGGTCAGAGANGTGAAGCTTGCACCAATTTGGGAAGTGGCCCACTGC 1268
Db 588 GCCAAGAACCCAGGTCAGAGATACAGAGCTTGCACCAATTTTGAAGCGGCTCGCTGC 647
QY 1269 CATTTTGGTAGC 1280
Db 648 CATCTTGCAGTC 659

```

```

RESULT 26
AG099717
LOCUS
DEFINITION Pan troglodytes DNA, clone: PTB-102B07.F, genomic survey sequence.
ACCESSION
AG099717
VERSION
AG099717.1 GI:16720234
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Tocoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2
(bases 1 to 712)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Tocoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111 Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13

```

```

LIBRARY
Vector : pKSI45
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..712
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-102B07.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
ORIGIN
Query Match 34.2%; Score 455; DB 9; Length 712;
Best Local Similarity 86.2%; Pred. No. 2.6e-126; Indels 9; Gaps 2;
Matches 529; Conservative 0; Mismatches 76;

QY 671 GAGCTCACTAAAATGCTAATCAGGCAAAAAACAGGAGGTAAAGCAATAGCCAAATCATCTAT 730
Db 107 GAGCTCACTAAAATGCTAATTA-GCAAAAACAGGAGGTAAAGAAATAGCCAAATCATCTAT 165
QY 731 TGCCTGAGAGCAGCGGGAAGGACAGGATTTGGGATATAAACTCAGGCAATTCAGGCCAG 790
Db 166 TGCCTGAGAGCATAGCGGAGGAGGACAAATGATCGGATATAAAACCCAGGCAATTCAGGCCAG 225
QY 791 CACAGCAACCCCTTTGGGTCCCTCCATTTGATGGGAGCTCTGTTTTCACCTCTATTT 850
Db 226 CAAGGCTACCTCTTTGGGTCCCTCTTTGATGGGAGCTCTGTTTTCACCTCTATTT- 284
QY 851 CACTCTAATTAATCATGCAACTGCACTCTTCTGGTCCGTGTTTATGGCTCAAGCTGA 910
Db 285 -----ATATCTCGCAACTGCACTCTCTTGGCCTGTATTTGTTATGGCTCTAGCTGA 337
QY 911 GCTTTTGTTCGCCATCCACACTGCTGTTTGCACCGTCACAGACCGGCTGCTGACTTCC 970
Db 338 GCTTTTGTGAGCGTCCACCACTGCTGTTTGTGCTGCCATCACAGACCGGCTGCTGACTTCC 397
QY 971 ATCCCTTTCGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCGAGGTACCCATTGCC 1030
Db 398 ATCCCTTTCGATCCGCGAGGTGTCTCTGTGCTCTGATCCAGTGAGATACCCATTGCC 457
QY 1031 ACTCCCGATCAGGCTAAAGCTTTGCCATTGTTCTTCATGGCTGAGTGCCTGGGTTGTC 1090
Db 458 ACTCCCGATTCGGCTAAAGACTTTGCCATTGTTCCCGCATGGCTAAGTGCACGCTCATC 517
QY 1091 CTAAATGAACTGAACACTGCTGCTCACTGGGTTCCATGGTTCTTTCATGACCCAGGCTTC 1150
Db 518 CTAAATCGAGCTGAACACTAGTCACTGGGTTCCACAGTTCTCTTCCATGATCCATGGCTTC 577
QY 1151 TAATAGCTATTAACACTCACCAGCATGGCCCAAGATTCATTCCTTGGTATCTGTGAGGC 1210
Db 578 TAATAGCTATTAACACTTACTCATGGCCGAAGATCTCAATCTTGGAACTGTGAGGC 637
QY 1211 CAAGAACCCCGAGGTGAGAGANGTGAAGGCTTGCACCAATTTGGGAAGTGGCCCACTGCCA 1270
Db 638 CAAGAACCCCGAGGTGAGAGACAGGAGGCTTGCACCAATTTTGGAAATGSCCCACCACCA 697
QY 1271 TTTTGTGAGCGGC 1284
Db 698 TCTTGAAGAGGCGC 711

RESULT 27
AG121669/c
LOCUS
DEFINITION Pan troglodytes DNA, clone: PTB-130M15.F, genomic survey sequence.
ACCESSION
AG121669
VERSION
AG121669.1 GI:16650834
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 689)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@sc.riken.go.jp, URL:http://hsp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..689
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-130M15.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
Query Match 33.0%; Score 439.2; DB 9; Length 689;
Best Local Similarity 87.0%; Pred. No. 1.6e-121;
Matches 507; Conservative 0; Mismatches 66; Indels 10; Gaps 2;
Qy 257 CCCTGTATCTTCAACTTCTTGTAAAGTTTCTCTCTCCAGAAATTGAAGCTGTAAGCTA 316
Db 666 CCNTGGATCTTTAAAGTCCTTTTAAAGCTTGTCTCTCCAGAAATCGAAGCTGTAACCA 607
Qy 317 CAATATAGTTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAAATCTACCGTGGACCC 376
Db 606 TAATAGTTTCTTCAAATGGAGACCTTATGCAGTCCATGACTAAGATCTACCATGGACCC 547
Qy 377 CTGACCGGCTGTAGACTATGCTGTATGATGATTAATGACATTGAAGTCACCCCTCCCGAG 436
Db 546 CTGACCAAGCTGTAGCCCATGCTCTGATGTTATGACATCGAAGGACCCCTCTCTGAG 487
Qy 437 GAAATCTCAACTGCAC -AACCCCTACTACCTCCAAATTCACTAGTAGAAGCAGTTAGACAG 495
Db 486 GAAATCTCAACTGCACAAACCTTACTACATCCAGTTCAAGGAAAGCAGTTAGAGCGG 427
Qy 496 TTGTCAGGCAACCTCCCAACAGTACTTGGGTTTTCTCTTTGAGAGGGTGCAGCTAGAGA 555
Db 426 TCGTCGGCAACCTCCCAATGGCACTTGGGTTTTCTCTTTGAGAGGGGGGACTAGAGA 367
Qy 556 CAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCCTANTCTGGGAAGGTGAC 615
Db 366 CAGNACTATCTGGATTTCTTAGGCCAACTAAGAAATCCCTAAGCCTAGCTTGGGAAGGTGAC 307
Qy 616 CGCATCCATCTTTAAACATGGGGCTTTGCAACTTAGCTCACACCCGACCAATC----- 667
Db 306 CACATTCTCTTTAAACACAGGGCTTGAACCTTAGCTCACACCCGACCAATCAGTAGTA 247
Qy 668 -AGAGAGCTCTACTAAATAGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCCAATCAT 726
Db 246 AAGAGGGCTCACTAAATAGCTTAATTAGGCNAAAACAGGAGGTAAAGAAATAGCCAAATAT 187
Qy 727 CTATTTCCCTGAGACACAGCGGGAAGGACAAAGGATTGGATATAAATCTCAGCAATTCAG 786

Db	186	TTATTGCTCCAGAGTACGCGGAGGACAAATGATCGGATATAAACTATGATTCGAG	127
Qy	787	CCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGG	829
Db	126	CCACCAATGCTACCTCTTTGGGTCCCTCCCTTTGTTATGG	84

RESULT 28	BH149565/c
LOCUS	ENTQ487R Entamoeba histolytica Sheared DNA linear GSS 27-AUG-2001
DEFINITION	genomic, genomic survey sequence.
ACCESSION	BH149565
VERSION	BH149565.1 GI:15310303
KEYWORDS	GSS.
SOURCE	Entamoeba histolytica
ORGANISM	Entamoeba histolytica
REFERENCE	Eukaryota; Entamoebidae; Entamoeba.
AUTHORS	1 (bases 1 to 921) Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.
TITLE	Determination of clone end sequences from Entamoeba histolytica
JOURNAL	HM1:IMSS sheared DNA library (2001)
COMMENT	Unpublished (2001) Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: biloftus@tigr.org Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library Seq primer: M13-Reverse Class: shotgun High quality sequence start: 40 High quality sequence stop: 567. Location/Qualifiers 1. .921 /organism="Entamoeba histolytica" /mol_type="genomic DNA" /strain="HM1:IMSS" /db_xref="taxon:5759" /clone_lib="Entamoeba histolytica Sheared DNA" /note="Vector: pHOSt; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, a C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:1450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell. Oxford University Press, 1999)."

ORIGIN									
	Query Match	33.0%	Score 438.8	DB 8	Length 921				
	Best Local Similarity	83.8%	Pred. No. 2.4e-121						
	Matches 550	Conservative	0	Mismatches 84	Indels 22	Gaps 4			
Qy	547	ACTGAGACAGGACTAGCTGGATTCTCTAGGCTGCTAAGAAATCCCNAAAGCCTTANCTGG	606						
Dd	651	AGTGAGACAGGACTAGCTGGATATCTTAGGCGCTACTAAGAAATCCTTTAAGCTTAGCTGG	592						
Qy	607	GAAGGTGACCGCATCCCATCTTTTAAACATGGGGCTTGCNACTAGCTCACACCCGACCAAT	666						
Dd	591	GAAGGTGACTGCATCCACCTATAAACACGGGGCTTGGCAACTTAGCTCACACCCGACCAAT	532						
Qy	667	C-----AGAGAGCTCACTAAATGTCTAATCAGGCATAAAACAGGAGGTAAGACCAATA	717						


```
Db 531 CAGGTAGTAAGAGAGCTCACTAAACCGCTAATAGGCAAAACACAGGAGTTAAAGAAATA 472
Qy 718 GCATATCATCTATTGCTGAGAGCAGCGGAGACAGAGATGGGATATAAACTCAG 777
Db 471 GCATATCATCTATTGCTGAGAGCAGCGGAGAGCAATGA-TAGGATATAAAACCCAG 413
Qy 778 GCATTCACGCGCAGCAACACACCCCTTTGGGTCCCTCCCATCTGATGGAGCTCTGT 837
Db 412 GCATTCACGCGCAGCAACAGGCTACCTCTTTGAGTCCCTCCCTTTGATGGAGCTCTGT 353
Qy 838 TTTCACTCTATTTCATCTATTAAATCATGCAACTG--CACTCTCTGCTCGTGTGTTT 895
Db 352 -----TTTCACTCTATTAAATCTTGCAACTCTGCACACTCTCTGCTCAGTGTGTT 303
Qy 896 TATGGCTCAAGCTGAGCTTTGTTCCCATCCACCACTGCTGTTCCACCGTCAAGAC 955
Db 302 TAGCACTCGAGCTGAGCTTTCACTTGGCGTCCACCACTGCTGTTTCCCGCATCGAGAC 243
Qy 956 CGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTGATCCAGC 1015
Db 242 CGCGCGCTGACTTCCACCTCCAGATCTGGCAGGGTGTCTGCTGATCCTGATCCAGT 183
Qy 1016 GAGGTACCATGCTCCACTCCGATCAGGCTAAAGGCTTGGCATTTCTGATGGCTAA 1075
Db 182 GAGGCACCATTTGCGCTCTGGATTGGGCTAAAGTCTTGTCAATTTCTTGCACGGCTAA 123
Qy 1076 GTGCTGGGTTGCTCTAATAGACTGAACACTGGTCACTGGTTCATGGTTCCTCTCC 1135
Db 122 GTGCGCGGTTCACTTAATAGAGCTGAACATTTAGTCTGGTGGTTCACAGTTCTCTCT 63
Qy 1136 ATGACCCACCGCTTCTAATAGAGCTATAACACTCACCGCATGCCCCAAGATTCAT 1191
Db 62 GAGACCCACCGCTTCTAATAGAGCTATAACAGCACTGATGCCCCAAGATTCAT 7
```

```
RESULT 29
AG076758/c
LOCUS AG076758 679 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-071A23.R, genomic survey sequence.
ACCESSION AG076758
VERSION AG076758.1 GI:16628560
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 679)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpbes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .679
/organism="Pan troglodytes"
/mol_type="genomic DNA"
```

```

/db_xref="taxon:9598"
/clone="PTB-071A23.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 32.8%; Score 435.6; DB 9; Length 679;
Best Local Similarity 86.3%; Pred. No. 2e-120;
Matches 480; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 133 AATATAATGTAGAGAGAGGACCTTCAAAACACTGCACCTCGGGCCTCCTCAGCCAATG 192
Db 599 AACTCAATGTAGAGCATAGTAGTTTCAGAGCCCGGTCAATGGGGCTCGTCAGTCAATG 540
Qy 193 GATGCCCTGGAGCTCTCCCTCTTAGGACCTCTAGCAGCTATAATAATATTTTACTCTCTT 252
Db 539 GATGCCCTGGGTTCGCCCTCTTAGACCTCTAGCAGCTGTAGAGCTGTAGGCTGTGCTCT 480
Qy 253 TGGACCCCTGTATCTTCAACTCTCTTTAAAGTTTGTCTTCCAGAAATTGAAGCTGTAAA 312
Db 479 TGGGCCCTGTATATTTCGACCTCTTGTGAGTTTGTCTCTTCAGAAATTAAGGCTGTAAA 420
Qy 313 GCTACAAATAGTTCTTCAAAATGGAACCCAGATCGATCCATCACTAAAATCTACCGTGG 372
Db 419 GCTACAAATAGTTCTTCAAAATGGAACCCAGATCGATCCATCACTAAAATCTACCATGT 360
Qy 373 ACCCTGTGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGAAGTCAACCCCTCC 432
Db 359 ACCCTGTGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGAAGTCAACCCCTCC 300
Qy 433 CGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGGAAGCTAGAG 492
Db 299 TGGGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGGAAGCTAGAG 240
Qy 493 CAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGTTGAGCTGAG 552
Db 239 CGTCTGCTCAGTCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGTTGAGCTGAG 180
Qy 553 AGACAGGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCNAAGCTTANTCTGGGAAGGT 612
Db 179 AGACAGGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCNAAGCTTANTCTGGGAAGGT 120
Qy 613 GACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCGCAATACAGAGA 672
Db 119 GACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCGCAATACAGAGA 60
Qy 673 GCTCACTAAAATGCTTA 688
Db 59 GCTCTCGAGCATGCTA 44

RESULT 30
AG102951
LOCUS AG102951 683 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-106G16.F, genomic survey sequence.
ACCESSION AG102951
VERSION AG102951.1 GI:16723468
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 683)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
```

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbee@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .683

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-106G16.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 32.6%; Score 433.6; DB 9; Length 683;

Best Local Similarity 88.6%; Pred. No. 8.1e-120;

Matches 514; Conservative 0; Mismatches 61; Indels 5; Gaps 4;

QY 547 ACTGAGACGAGCTAGCTGGATTCTTCTAGCTGACTAAGAATCCCNAGCCTTANCTGG 606

Db 108 ATTGAGAGAGGACTAGCTGGATTCTTCTAGCCGACTAAGAATCCCTAAGCCTAGCTGG 167

QY 607 GAAGTGACCGCATCATCTTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAAT 666

Db 168 GAAGTGACTGCATCCACCTTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAAT 227

QY 667 CAGAGGCTCACTAAATGCTAATCAGCAAAACAGGAGTAAAGCAATAGCCCAATCAT 726

Db 228 CAGAGGCTCACTAAATGCTAATTAGGCAAAACAGGA-GTAAAGAAACAGCCCAATCAT 286

QY 727 CTATTGCTCTGAGACACAGCGGGAAGGCAAGGATTGGGATATAAATCAGGCATTCAG 786

Db 287 CTATTGCTCTGAGACACAGCGGGAAGGCAAGGATCGGATATAAATCAGGCATTCAG 346

QY 787 CAGACAGACAGACCCCTTTGGTCCCTCCCATCTGATGGAGCTCTGTTTCACTCT 846

Db 347 CCGGCAACGCAACCCCTTTGGTCCCTCCCTCC--TTGTAAGGAATCTGTTTCACTCT 404

QY 847 ATTTCACTCTATTAAATCATGCACTGCACTCTCTGCTCCCTGTTTATGGCTCAAG 906

Db 405 ATTTCACTCTATTAAATCTGCAACTGTGCTCTTCTGCTCCCTGTTTATGGCTCAAG 464

QY 907 CTGAGCTTTTGTTCGCCATCCACCATCTGTG--TTTGCCACCGTCAAGACCCGCTGTGA 965

Db 465 CTGAGCTTTTCACTTGCTGCTCCACCATCTGTGTTTGTGCTCCCTGTTTATGGCTCAAG 524

QY 966 CTTCCATCCCTTTGGATCAGACAGAGTCCACTGTCTCTGATCCAGGAGGTACCCA 1025

Db 525 CTTCCATCCCTCCGGATCAGACAGAGTGTGGGCTGGGCTCCCTGATCCAGTGAGACCCCA 584

QY 1026 TTGCCACTCCGATCAGGCTAAGGCTTGCCATTGTTCTGATGGCTAAGTGCCTGGGT 1085

Db 585 TTGCC--GTTCCGATTGGGCTAAGGCTTGCCATTGTTCTGATGGCTAAGTGCCTGGAT 643

QY 1086 TTGTCTTAATAGAACTGGAACACTGGTCACTGGGTTCATG 1125

Db 644 TCATCCTAATCGAGCTGAACACTAGTCACTGTGTTCAGG 683

RESULT 31

AQ381711

LOCUS

DEFINITION AQ381711 Homo sapiens genomic clone RPCI-11-165G20, GSS 21-MAY-1999

genomic survey sequence.

AQ381711

AQ381711.1 GI:4352734

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 611)

Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and

Venter,J.C.

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

Other GSSs: RPCI11-165G20.TV

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@igr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genetics (info@resgen.com). BAC end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .611

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="GDB:7563139"

/db_xref="taxon:9606"

/clones="RPCI-11-165G20"

/sex="Male"

/cell_type="Lymphocytes"

/clone_lib="RPCI-11"

/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

RPCI11 Human Male BAC Library"

ORIGIN

Query Match

Best Local Similarity

Matches 517; Conservative

0; Mismatches 62; Indels 21; Gaps 3;

Score 430.6; DB 8; Length 611;

Pred. No. 6.4e-119;

QY 597 GAATCCCNAAAGCTTANCTGGGAAGTGACCGCATCCATCTTTAAACATGGGGCTTGCAAC 646

Db 22 GAATTCCTAAACCTAGCTGAGAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAAC 81

QY 647 TTAGCTCACACCCGACCAATC-----AGAGAGCTCACTAAATCTTAATCAGGCAA 697

Db 82 TTAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCTCACTAAATGCTTAATAGGCAA 141

QY 698 AAACAGGAGGTAAAGCAATAGCCAATCATCTATTTCCTGAGAGCACAGCGGGAAGGCAA 757

Db 142 AAACAGGTGGTAAAGAAACAGCCAATCATCTATTTCCTGAGAGCACAGTGGGAGGCAA 201

QY 758 GGATGGGATATAAATCAGGCATTCAGCAGCAAGCAACCCCTTTGGTCCCTC 817

Db 202 TGATCGGATATAAATCCAGGCATTCGAGCCAGCAATGGCTACCTCTTTGGGTCCCTC 261

QY 818 CCATGTATGGAGCTGTGTTTCACTCTATTTCCTGATCTATTAAATCATGCAACTG--CA 875

Db 262 TCTTTGATGGAGCTCTGT-----TTTCACTCTATTAAATCTTGACCTGTACT 311

QY 876 CTCCTTCTGGTCCGTGTTTTTATGGCTCAAGCTGAGCTTTGTCGCCATCCACCACTGC 935

Db 312 CTCCTTCTGGTCCGTGTTTTTATGGCTCAAACTGAGCTTTCACTCGCCATCCCAACTGC 371

QY 936 TGTTGGCCACCGTCAACACCCGCTGTGACTTCCATCCCTTTGGATCCAGCAGATGTC 995


```
/mol_type="genomic DNA"
/db_xref="GB:7524474"
/db_xref="taxon:9606"
/clone="RPCI-11-64L19"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/notes="Vector: pBCE3.6; Site 1: EcoRI; Site 2: EcoRI;
RPC111 Human Male BAC Library"

ORIGIN
Query Match      32.2%; Score 427.6; DB 8; Length 714;
Best Local Similarity 83.0%; Pred. No. 5.1e-118;
Matches 565; Conservative 0; Mismatches 92; Indels 24; Gaps 6;

QY 589 ATCCNNAAGCCTTANT-GGGAAGGTGACCGATCATCTTTAAACATGGGCTTGCACACT 647
Db 1 ATTCCTAAGCCTAGCTGGGAAGGTGACCAACACCCCTCTTTAAACACAGAGCTTGTACT 60

QY 648 TAGCTCACACCGGACCAATC-----AGAGGCTCACTAAATGCTAATCAGGC-AA 697
Db 61 CAGCTCACACCGGACCTATCAGGTAGTAAGAGAGGCTCACTAAATACCAATTAGGCTAA 120

QY 698 AAACAGGAGGTAAAGCAAT-AGCCAAATCATCTATTGCTTGAGAGCACAGCGGGAAGGACA 756
Db 121 AAACAGGAGGTAAAGAAATATCAATCATCTATCGCTTGAGAGCACAGGGGGAGGACA 180

QY 757 AGGATTGGGATATAAATCAGGATTCAGCCAGTCAAGCCAGCAGACAGCAACCCCTTTGGTCCCT 816
Db 181 ATGATCGGGATATAAACCAGGCAATTTGAGCCAGATCAGGTAAACCTCTTTGGGTCCCT 240

QY 817 CCCATTGTATGGAGCTGTGTTTTCACCTCTATTTCACCTCTAATTAATCATGCACTG--C 874
Db 241 CACACTGTATGGAGCTCTGT-----TTTCACTCTAATAATCTTGCACTGAC 290

QY 875 ACTCTTCTGGTCCGCTGTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCCACTG 934
Db 291 ACTCTTCTGGTCCATGTTGTTCCGCTCAAGCTGAGCTTTTGTCTCGCGCTCCACCACTG 350

QY 935 CTGTTTGGCACCGTCACAGACCCGCTGTGACTTCGATCCCTTTGGATCCAGCAGAGTGT 994
Db 351 CTGAATGGCGCATTCGAGACCTGCGCTTGACTTCCACCCCTCCGGATCCGGCAGAGTGT 410

QY 995 CCACGTGTCTCTGATCCAGCAGGTACCCATTGCGACTCCCGATCAGGCTAAAGGCTTG 1054
Db 411 CCGCTGCACTCTGATCCAGCAGGACCCATTGCGACTCCCGATCAGGCTAAAGGCTTG 470

QY 1055 CCATTGTTCTCGATGGCTAAGTGGCTTGTCTTAATAGAACTGAACACTGGTCAAC 1114
Db 471 CCATTGTTCTCGACAGCTAAGTGGCTTGTCTTAATAGAACTGAACACTGGTCAAC 530

QY 1115 TGGGTTCCATGTTCTCTTCATGACCCACGGCTTCTAATAGAGCTAACAACCTCACCGC 1174
Db 531 TGGGTTCCACGGTTCTCTTCCATGACTCACAGCTTCTAATAGAGCTAACAACCTTACCAC 590

QY 1175 ATGGCCCAAGATTCCATTCTTTGGTATCTGTGAGCCAAAGACCCAGCTCAGAGAANGT 1234
Db 591 ATGGCCCAAGTTTCAATTCGTTGGATCCATATGCCAAGAACCACCGGTGAGAGAATAA 650

QY 1235 GAGGCTTGCCACCATTTGGGA 1255
Db 651 AAGGCCCGCCCATCTGGGA 671

RESULT 34
AG108186
LOCUS AG108186 697 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-113D17.R, genomic survey sequence.
ACCESSION AG108186
VERSION AG108186.1 GI:16728704
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 697)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chiebes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI.
R.Site 2 : SacI.
Location/Qualifiers
1. .697
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-113D17.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      31.9%; Score 424.4; DB 9; Length 697;
Best Local Similarity 82.0%; Pred. No. 5.1e-117;
Matches 488; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 669 GAGAGCTCATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCT 728
Db 95 GAGAGCTCATAAATGCTAATTAGGCAAAACAGTAGGTAAAGAAATAGCAATCATCT 154

QY 729 ATTCCCTGAGACACAGCGGAGGACAGGATGGGATATAAACTCAGGCATTCAAGCC 788
Db 155 ATTCCCTGAGACACAGCGGAGGACAGGATCCGGATATAAACCCAGGCATTCAAGCT 214

QY 789 AGCAACAGCAACCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTCACTCTAT 848
Db 215 GGCAACGAAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTCACTCTAT 274

QY 849 TTCACTCTAATAATCATGCACTGCTCTTCTGGTCCGTTTATGGCTCAAGCT 908
Db 275 TTCACTCTAATAATCATGCACTGCTCTTCTGGTCCGTTTATGGCTCAAGCT 334

QY 909 GAGCTTTTGTTCGGCATCCACCACTGCTGTTTTCACCGCTCACAGACCCGCTGCTGACTT 968
Db 335 GAGCTTTTGTTCGGCATCCACCACTGCTGTTTTCACCGCTCACAGACCCGCTGCTGACTT 394

QY 969 CCATCCCTTTGGATCCAGCAGAGGTGCTCATGTGCTCTCTGATCCAGCAGGTACCCATTG 1028
Db 395 ACATCCCTTTGGATCCAGCAGAGGTGCTCATGTGCTCTCTGATCCAGCAGGTACCCATTG 454

QY 1029 CCATCCCGATCAGGCTAAAGGCTTGCCATTGTTTCTGTCATGGCTAAGTGCCTGGGTTG 1088
Db 455 CCGTTCCTCTATCTGGAATAAAGCCCGCATTTGCTTCCGACGGCTAACCGCTCGGGTTCA 514

QY 1089 TCCTAATAGAACTGAACATGCTGGTTCACCTGGTTCACATGTTCTCTTCCATGACCCAGGCT 1148
Db 515 TCCTAAGCAGAGTGCACCACTAGTCACTGGGTCCCTGCGTCTCTCTCATGACCCAGGCT 574

QY 1149 TCTAATAGAGTATAAACAACCTCACCGCATGGCCCAAGATTCCATTTCCTTGGTATCTGTGAG 1208
```


Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."		High quality sequence stop: 552.	
FEATURES	source	1..803	Location/Qualifiers
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:30337259"	
		/tissue_type="pre-eclamptic placenta"	
ORIGIN		/lab_host="DH10B Tona"	
		/clone_lib="NIH_MGC_148"	
		/notes="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."	
		Query Match 31.8%; Score 422.2; DB 6; Length 803;	
		Best Local Similarity 88.5%; Pred. No. 2.5e-116;	
		Matches 469; Conservative 0; Mismatches 59; Indels 2; Gaps 1;	
ORIGIN		QY 800 CCCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTCACTCTATTTCACCTATT 859	
		DB 34 CCCCCTTTGGGTCCCTCCCTTTGATGGAG--CTGTTTTCATGCTATTTCACCTATT 91	
ORIGIN		QY 860 AAATCATGCAACTGCATCTTTCGGTCCGTGTTTTTATGGCTCAAGCTGAGCTTTTGT 919	
		DB 92 AAATCTTGAACCTGCATCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTGCT 151	
ORIGIN		QY 920 CGCCATCCACCACTGCTTTGCCACCGTTCACAGACCGCTGCTGACCTCCATCCCTTTG 979	
		DB 152 CACCGTCCACCACTGCTTTGCCACCGTTCACAGACCGCTGCTGACCTCCATCCCTCTG 211	
ORIGIN		QY 980 GATCAGCAGAGTCTCACTGTGCTCTGATCCAGCAGGTACCCATTGGCCATCCCGAT 1039	
		DB 212 GATCTGCGAGGTGTCCGCTGTCTCTGATCCAGCAGGCGCCATTGGCGCTCCCAAT 271	
ORIGIN		QY 1040 CAGGCTAAAGGCTTGGCCATTGTTCTGTCATGGCTAAAGTGCCTGGGTTGCTTAATAGAA 1099	
		DB 272 TGGGCTAAAGGCTTGGCCATTGTTCTGTCATGGCTAAAGTGCCTGGGTTTGTCTAATTGAG 331	
ORIGIN		QY 1100 CTGAACACTGGTCACTGGGTTCCATGTTCTTCTCAATGACCCAGGCTTCTAATAGAGC 1159	
		DB 332 CTGAACACTAGTCACTGGGTTCCATGGTTCCTTCTGTGACCCACGGCTTCTAATAGAAC 391	
ORIGIN		QY 1160 TATAACACTCACCGATGCCCAAGATTCCATTCTTGGTATCTGTGAGGCCAAGAACCC 1219	
		DB 392 TATAACACTTACCACTGCGCCAGATTCATCTTGGAAATCGTGAGGCCAAGAACTC 451	
ORIGIN		QY 1220 CAGGTCAGAGANGTGGAGCTTGGCCACCATTTGGAAAGTGGCCCACTGCCATTTTGGTAG 1279	
		DB 452 CAGGTCAGAGAAATACAGGCTTGGCCACCATCTTGGAAAGCGGCTGTACCATCTTGGAG 511	
ORIGIN		QY 1280 CGGCCCAACCACTTTGGGAGCTGTGGAGCAAGGATCCCCCAGTAACA 1329	
		DB 512 TGGTTTCAACCACTTTGGGAGCTCTGTGAGCAAGGACCCCGGTAACA 561	
RESULT 37			
CB992422			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 826)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM359 row: k column: 16
High quality sequence stop: 585.

FEATURES

1..826
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3036351"
/tissue_type="pre-eclampsia placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: Placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to 10⁵. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 31.8%; Score 422.2; DB 6; Length 826;
Best Local Similarity 88.5%; Pred. No. 2.5e-116;
Matches 469; Conservative 0; Mismatches 59; Indels 2; Gaps 1;
QY 800 CCCCCTTTGGTCCCTCCCATGTATGGAGCTCTGTTTCACTCTATTCTACTCTATT 859
DB 34 CCCCCTTTGGTCCCTCCCTTTGTATGGAG--CTGTTTTCATGCTATTCTACTCTATT 91
QY 860 AAATCATGCAACTGCACTCTCTGTCGCGTGTGTTTATGGCTCAAGCTGAGCTTTGTT 919
DB 92 AAATCTTGGCACTGCACTCTCTGTCGCGTGTGTTTATGGCTCGAGCTGAGCTTTGCT 151
QY 920 CGGCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCTGCTGACTTCCATCCCTTTG 979
DB 152 CACGCTCCACCACTGCTGTTTGGCCACCGCAGACCTGCGCTGACTCCATCCCTCTG 211
QY 980 GATCCAGAGAGTGTCCACTGTGCTCTCTGATCCAGCGAGTACCCATGCGACTCCCGAT 1039
DB 212 GATCCTGAGGAGTGTGCGCTGTGCTCTGATCCAGCGAGGCGCCATGCGCGTCCCAAT 271
QY 1040 CAGGCTAAAGCTTGCCATGTTCTCTGATGGCTTAAGTGTGCTGGTGTGCTTAATAGAA 1099
DB 272 TGGGCTAAAGCTTGCCATGTTCTCTGATGGCTTAAGTGTGCTGGTGTGCTTAATAGAG 331
QY 1100 CTGAACACTGCTGCTGGTGTGCTGTTCTCTTCCATGACCCACCGCTTCTAATAGAGC 1159
DB 332 CTGAACACTGCTGCTGGTGTGCTGTTCTCTTCTGATGCCACCGCTTCTAATAGAAC 391
QY 1160 TATAAACAATCCGATGGCCCAAGATTCATTCCTTGGTATCTGTGAGGCGCAAGACCC 1219
DB 392 TATAAACAATCCGATGGCCCAAGATTCATTCCTTGGTATCTGTGAGGCGCAAGACCC 451
QY 1220 CAGGTACAGAGTGTGCTGCTGCTGATCCAGCGAGGTACCCATGCGACTCCCGAT 1279

DB 452 CAGGTACAGAGTGTGCTGCTGATCCAGCGAGGTACCCATGCTGCTACTCTTTGGAAG 511
QY 1280 CGGCCCCACCACTCTTGGAGCTCTGGAGCAAGGATCCCCAGTAACA 1329
DB 512 TGGTTCAACCACTCTTGGAGCTCTGTGAGCAAGGACCCCGGTAAACA 561

RESULT 39

CB996320
LOCUS
DEFINITION
AGENCOURT 13617032 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30332232 5', mRNA sequence.

ACCESSION

CB996320
VERSION
CB996320.1 GI:30290840

KEYWORDS

SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 837)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov

AUTHORS

Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM348 row: p column: 01
High quality sequence stop: 611.

FEATURES

Location/Qualifiers
1..837
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30332232"
/tissue_type="pre-eclampsia placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: Placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to 10⁵. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 31.8%; Score 422.2; DB 6; Length 837;
Best Local Similarity 88.5%; Pred. No. 2.5e-116;
Matches 469; Conservative 0; Mismatches 59; Indels 2; Gaps 1;
QY 800 CCCCCTTTGGTCCCTCCCATGTATGGAGCTCTGTTTCACTCTATTCTACTCTATT 859
DB 34 CCCCCTTTGGTCCCTCTCCCTTTGTATGGAG--CTGTTTTCATGCTATTCTACTCTATT 91
QY 860 AAATCATGCAACTGCACTCTCTGTCGCGTGTGTTTATGGCTCAAGCTGAGCTTTGTT 919
DB 92 AAATCTTGGCACTGCACTCTCTGTCGCGTGTGTTTATGGCTCGAGCTGAGCTTTGCT 151
QY 920 CGGCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCTGCTGACTTCCATCCCTTTG 979
DB 152 CACGCTCCACCACTGCTGTTTGGCCACCGCAGACCTGCGCTGACTCCATCCCTCTG 211
QY 980 GATCCAGAGAGTGTCCACTGTGCTCTGATCCAGCGAGGTACCCATGCGACTCCCGAT 1039

Db	212	GATCCTCGAGGGTGTCCGGCTGTGCTCTGATCAGCGAGGCGCCATTCGCGCTCCCAAT	271
Qy	1040	CAGGCTAAAGGCTTGCCATTGTTCTCGATGGCTTAAGTCGCTGGTGTGCTCTAAATAGAA	1099
Db	272	TGGGCTAAAGGCTTGCCATTGTTCTCGACGGCTAAGTCGCTGGTGTGCTCTAAATGAG	331
Qy	1100	CTGAACACTGGTCACATGGGTTCCATGGTTCCTTCCATGACCCACGGCTTCTAATAGAGC	1159
Db	332	CTGAACACTAGTCACATGGGTTCCATGGTTCCTTCTGTGACCCACGGCTTCTAATAGAAC	391
Qy	1160	TATAACACTCACCGCATGGGCCCAAGATTCATTCCTTGGTATCTGTGAGGCCAAGACCC	1219
Db	392	TATAACACTTACCACATGGGCCAAGATTCATTCCTTGGAAATCCGTGAGGCCAAGACTC	451
Qy	1220	CAGTCTCAGAAAGTGAAGGCTTGCCACCATTTGGGAAGTGGCCACCTGCCATTTTGGTAG	1279
Db	452	CAGTCTCAGAAATACGAGGCTTGCCACATCTTGGAAAGCGGCTGTACCATCTTGGAG	511
Qy	1280	CGGCCCAACCACTCTTTGGGAGCTGTGGGAGCAAGTATCCCCCAAGTAACA	1329
Db	512	TGGTTTACCAACCATCTTTGGAGCTCTGTGACGAAGACCCCGCGGTAAACA	561

RESULT 40
 CB994833
 LOCUS
 DEFINITION
 CB994833
 IMAGE:303379155 5', mRNA sequence.
 VERSION
 KEYWORDS
 EST.
 CB994833.1 GI:30289353
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 852)
 NH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Straubeberg, Ph.D.

FEATURES
SOURCE

ORIGIN